FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

UNITED NATIONS ENVIRONMENT PROGRAMME

MANUAL OF METHODS IN AQUATIC ENVIRONMENT RESEARCH
Part 5 - Statistical Tests

by

F. Möller
Istituto di Economia e Politica Agraria
University of Milan, Italy

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
Rome, 1979

The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

M-45
ISBN 92-5-100767-5

The copyright in this book is vested in the Food and Agriculture Organization of the United Nations. The book may not be reproduced, in whole or in part, by any method or process, without written permission from the copyright holder. Applications for such permission, with a statement of the purpose and extent of the reproduction desired, should be addressed to the Director, Publications Division, Food and Agriculture Organization of the United Nations, Via delle Terme di Caracalla, 00100 Rome, Italy.

© FAO 1979

PREPARATION OF THIS MANUAL

This manual was prepared as part of a cooperative project of the United Nations Environment Programme entitled

FAO(GFCM)/UNEP Joint Coordinated Project on Pollution in the Mediterranean

with the Food and Agriculture Organization of the United Nations as cooperating agency.

Distribution:

FAO Fisheries Department
FAO Regional Fishery Officers
Mediterranean Pollution Project
Selector SM/SI
UNEP

Bibliographic reference:

Möller, F. (1979)
FAO Fish.Tech.Pap., (182):131 p.
Manual of methods in aquatic environment research. Part 3. Statistical tests

FAO manuals. Water pollution.
Statistical analysis. Pollution surveys
Bioassays.

PREFACE

Under the Mediterranean Action Plan developed by the United Nations Environment Programme, the Co-ordinated Mediterranean Pollution Monitoring and Research Programme was established. As part of this programme, an FAO(GFCM)/UNEP Joint Co-ordinated Project on Pollution in the Mediterranean was initiated. The project aims at organizing baseline studies and monitoring of metals, particularly mercury, and of DDT, PCBs and other chlorinated hydrocarbons, in marine organisms, and at establishing research projects on the effects of pollutants on aquatic organisms and ecosystems.

Further to a Manual on Sampling and Analyses of Biological Material (FAO Fish.Tech. Pap. 158), this Manual on Statistical Tests has been prepared to assist laboratories participating in the FAO/UNEP project in proper planning of monitoring programmes and research projects, and to facilitate the evaluation of data and results obtained. However, although the examples in the manual were chosen to specifically suit the FAO/UNEP project, this manual is expected to be of use to all laboratories involved in analysing pollution levels and measuring effects of pollution on aquatic life.

The manual was prepared by Dr. F. Möller, Milano. Drs. M. Bernhard, Fiascherino, and W. Besch, Karlsruhe, assisted in the design of the examples given. Final editing was carried out by Dr. H. Naeve, Fishery Resources and Environment Division, FAO. The views expressed in this manual are those of the author and do not necessarily represent the views of either FAO or UNEP.

CONTENTS

		Page
1.	TESTS FOR POLLUTION MONITORING PROGRAMMES	1
	1.1 Introduction	1
	1.2 Population Parameters	1
	1.3 Simple Sample Statistics	3
	1.4 Composite Sample Statistics	4
	1.4.1 One-factor models	5
	1.4.2 Two-factor nested model	8
	1.5 Regression or Covariance Models	9
	1.6 The Problem of the Sample Size	11
	1.6.1 One sample	11
	1.6.2 Two samples	12
	1.7 Examples	14
	1.7.1 Example of a one-factor model	17
	1.7.2 Example of a regression test	18
	1.7.3 Example of a two-factor nested model with the determination of the variance components for individuals and pretreatment	19
	1.7.4 Example of a two-factor nested model with the determination of the variance components for samples and individuals	20
	1.7.5 Example of a covariance model	21
	1.7.6 Example of a comparison of three regression lines	23
	1.7.7 Example of a set-up of a (quality) control plan	27
	1.7.8 Example of the determination of the sample sizes in an experiment with two samples	29
2.	TESTS FOR BIOLOGICAL POLLUTION RESEARCH	31
	2.1 Introduction	31
	2.1.1 Hypotheses, tests and errors	31
	2.1.2 Confidence intervals and tests	35
	2.2 Hypotheses about Frequencies	36
	2.2.1 Test whether the frequency π of a character in a population equals π_0 or it is larger	36
	2.2.2 Test whether the frequency π of a character in a population equals π_0 or it is smaller	36
	2.2.3 Comparison of two frequencies: test of independence (both margins fixed)	37
	2.2.4 Comparison of c frequencies: test of homogeneity (one margin fixed)	38
	2.2.5 Comparison of an observed with a theoretical distribution: tests of goodness of fit	40
	2.2.6 Comparison of an observed with the normal distribution	41
	2.2.6.1 Probit analysis	41
	2.2.6.2 Normality tests	41
	2.2.7 Comparison of two distributions	45
	2.2.8 A combination of type I errors	48
	2.3 Hypotheses about Location Measures	48
	2.3.1 Test whether the mean μ, of an approximately normal population with unknown variance, equals a preassigned value μ_0	48
	2.3.2 Comparison of two means sampled from approximately normal populations with unknown variances	49
	2.3.3 Comparison of the location of two populations with any distribution	51
	2.3.4 Comparison of the means of c samples	51
	2.3.5 Multiple *a posteriori* comparison	53

		Page
2.4	Hypotheses about Variances	55
2.4.1	Test whether the variance σ^2 of an approximately normal population equals the preassigned value σ_0^2	55
2.4.2	Comparison of two variances of approximately normal populations	56
2.4.3	Comparison of c variances of approximately normal populations: tests of homoscedasticity	56
2.4.4	Comparison of c variances of any distribution	58
2.5	Hypotheses about the Interdependence (Correlation) and the Dependence (Regression) in Bivariate Variables	58
2.5.1	Test of no correlation in a bivariate normal population	58
2.5.2	Test whether the correlation coefficient ρ of a bivariate normal population equals a preassigned value ρ_0	59
2.5.3	Comparison of two correlation coefficients of bivariate normal populations	59
2.5.4	Tests of no correlation in bivariate populations with any distribution	60
2.5.5	Hypotheses about regression coefficients of normal bivariate populations	62
2.6	Other Tests	63
2.6.1	Test for outliers	63
2.6.2	Tests for randomness	64
2.7	Examples	66
2.7.1	Examples of tests about frequencies	66
2.7.2	Examples of tests about location measures	78
2.7.3	Examples of tests about variances	87
2.7.4	Examples of tests about regression and correlation	88
2.7.5	Examples of miscellaneous tests	90
REFERENCES		93
ANNEX:	STATISTICAL TABLES	99
A-1	Common logarithms of factorials	99
A-2	Correction factors for ranges c_n, d_n and g_n and critical values for the median h_n	99
A-3	Critical values of the normal variate z	100
A-4	Critical values of the t_ν variate	101
A-5	Critical values of the χ^2_ν variate	102
A-6	Critical values of the F_{ν_1,ν_2} variate	103
A-7	Critical values of the Kolmogoroff-Smirnov test	106
A-8	Critical values of the Mann-Whitney-Wilcoxon U-test	107
A-9	Acceptance intervals r_1, r_2 for the Run test r	108
A-10	Critical values of Wilcoxon's pair difference test	110
A-11	Critical values of the test $\lambda_n = (\bar{x}-\mu)/R$	111
A-12	Critical values of the studentized range $Q\nu,k$	112
A-13	Critical values of F_{max}	116
A-14	Critical values of Spearman's and Kendall's rank correlation coefficients	117
A-15	Critical values of Dixon's test for outliers r_n	117
A-16	Critical values of Grubb's test for outliers T_n	118

		Page
A-17	Critical values of $Z_n = R/s$ for outliers	119
A-18	Critical values of the mean successive difference test η_n^2 and of the serial correlation coefficient $R_{h \cdot n}$	120
A-19	The probit transformation $Y_{(x_j)}$	121
A-20	Minimum and maximum working probits, range and weighting coefficients	122
A-21	Critical values of Van der Waerden's X test	123
A-22	Critical lower values $L_{(\alpha)}$ of the binomial distribution	124
A-23	Coefficients a_i for the W test for normality	125
A-24	Critical values of the W test for normality	127
A-25	Critical values of the normal probability plot correlation coefficient r_n	128
INTERPOLATION		129

1. TESTS FOR POLLUTION MONITORING PROGRAMMES

1.1 Introduction

Statistical data are observations about characteristics or attributes of individuals (e.g. Hg concentration, weight of specimens). The group of individuals under study is called a (statistical) population. Such a population could be all *Mullus barbatus* of a given area. If the frequency distribution of the attributes of the population is known, it is possible to describe it by a density function (e.g. the normal distribution or distribution of Gauss), which is an analytical function defined by a number of parameters.

The entire population is often too large to be studied as a whole or perhaps can only be defined hypothetically. It is therefore more convenient to study a subset of the population. Such a subset is known as a sample. The goal of sampling is to get information about the frequency distribution of the population, or more precisely, about the population parameters. For this reason sample functions are built to estimate the population parameters. These estimates are also known as sample statistics or simply statistics. If the investigation is based on a sample, the estimates will depend on the particular sample chosen; a different value may be obtained, if a different sample is selected. This variation in the values from sample to sample is known as sampling variability. Its magnitude can be ascertained by the theory of probability. Generally, the difference between a statistic and the true population value, i.e. the sampling error, increases as more factors influence the statistic. Factors which influence the statistics may be the age of specimen, location and time of sampling, etc. The main task of the research worker is, therefore, to identify the possible factors, and the task of the statistician to establish the amplitude of the factors in terms of their variability or error. If an effective estimation of the frequency distribution is achieved, the form of the density function is known. If it is not, one has to decide on the form of the function. This problem is known as specification and does not have an easy solution. Sometimes it can be solved from theoretical considerations of how the frequency distribution is generated, but more often the distribution form is disregarded, concentrating the attention on a few parameters to which most distribution parameters are strictly related.

1.2 Population Parameters

A population is designated univariate if only one character (y) is considered, and multivariate if more than one character (x,y) is considered. The most common univariate parameters are the mean μ as a location measure, the variance σ^2, the standard deviation σ and the coefficient of variation $V = \frac{\sigma}{\mu} 100$ as dispersion measures.

The most important multivariate parameter is the covariance σ_{xy} as a measure of correlation.

For a finite population we could calculate the parameters if we had measurements for all the individuals (specimens) in the population. In practice this is generally impossible, certainly in the case of destructive sampling, so that we can only expect a true population. We define expected value of any quantity its average value, averaged over the entire population.

Even though the exact frequency distribution is unknown, the knowledge of the mean and the variance is sufficient to make precise statements about univariate frequency distributions with the use of Tchebycheff's inequality:

$$P\left\{ |y - \mu| < \lambda\sigma \right\} \geqslant 1 - \frac{1}{\lambda^2}$$

which gives us a lower limit $P\{\}$ to the amount of frequency lying within a range of $\pm \lambda\sigma$ on either side of the mean. If the population is unimodal, i.e. it has only one mode $(= \mu_0)$, which is the most frequent value in the distribution, then the inequality can be improved so that the lower limit to the amount of frequency lying within a range of $\pm \lambda\sigma_0$, $(\sigma_0^2 = \sigma^2 + (\mu - \mu_0)^2)$, on either side of the mode, is increased to $1 - (\frac{2}{3\lambda})^2$. This inequality is due to Camp-Meidell.

In Table I we compare the frequency area of a normal distribution within the intervals of $\pm \lambda\sigma$ with the above inequalities.

Table I
Frequency area of normal distributions

λ	Normal distribution	Inequalities of:	
		Tchebycheff for any distribution	Camp-Meidell for unimodal distributions
1	.683	$\geqslant 0$	$\geqslant$.556
2	.954	$\geqslant$.75	$\geqslant$.889
3	.997	$\geqslant$.889	$\geqslant$.951
4	>.999	$\geqslant$.938	$\geqslant$.972

It can be seen that estimates of the frequency by the inequalities of Tchebycheff and Camp-Meidell are fairly good for values of λ larger than 3. This means, if large differences between sample values or functions are observed, the conclusions are significant even if the frequency function is not known.

When two or more variables are involved the main interest lies in the joint distribution of the variables. We may describe this as the problem of statistical relationship for which a useful distinction may be made. We may be interested either in the interdependence between a number of the variables (not necessarily all), or in the dependence of one or more variables upon the others, which, and this is a fundamental difference, are often under control of the experimenter. This means that the levels of the dependent variable (species, time and area of sampling, etc.) can be chosen by the experimenter.

We can say that the study of interdependence leads to the theory of correlation and the study of dependence to the theory of regression.

1.3 Simple Sample Statistics

Characteristics, parameters or density distributions of the population can be estimated in several ways. The most common is to estimate the parameters through sample statistics. Of course the estimate might differ from the parameter in any particular case, and hence there is a margin of uncertainty. The extent of this uncertainty can be expressed in terms of the sampling variance of the estimator, which is the variance of the distribution function of the estimates, i.e. the frequency distribution of the estimates of all possible samples of the same size, which can be drawn from one population. In other words, the sampling variance is a measure of the precision of the estimates.

If we denote the population parameter by θ, then any interval which specifies two statistics, T_1 and T_2, such that the probability that the population parameter lies within the interval (T_1, T_2) is $1-\alpha$, where α is a preassigned positive constant (≤ 1), then the interval (T_1, T_2) is said to provide an interval estimate, or a confidence interval for the parameter θ at the confidence level $1-\alpha$. On the other hand, if L_1 and L_2 ($\geq L_1$) are two sample functions such that the probability that the interval (L_1, L_2) will include at least $100\gamma\%$ of the population, with γ a positive constant (≤ 1), is equal to $1-\alpha$, then the limits L_1 and L_2 are called the $100\gamma\%$ tolerance limits for the population at the confidence level of $1-\alpha$.

Whereas consistent estimates of parameters can generally be given without reference to the frequency distribution of the population, exact confidence and tolerance limits cannot be evaluated without knowing the population or parent distribution. In this event, distribution free methods can be used. If order characteristics of the population, like the median or the range (not parameters in the sense previously defined) are to be estimated, then non-parametric methods can solve the problem. It is even possible to estimate some parameters by non-parametric methods.

In the following we will confine ourselves to the estimation of problems which can arise with the monitoring project *. The population mean μ is estimated by the sample mean:

$$\bar{y} = \frac{1}{n} \sum y_j$$

hence, the sample mean is itself a statistic. The sampling variance of this statistic is measured by:

$$\sigma^2_{\bar{y}} = \frac{\sigma^2}{n} .$$

Its square root known as the standard error, $\sigma_{\bar{y}}$, should not be confused with the standard deviation σ.

The population variance σ^2 is estimated by the statistic (unbiased estimate)

$$s^2 = \frac{\Sigma(y_j - \bar{y})^2}{n-1} .$$

Examples for all discussed statistics can be found in Section 1.7

The numerator of the variance, $\sum(y_j - \bar{y})^2$, is referred to as the sums of squares (SS), and the ratio of SS with the degrees of freedom (DF = n - 1), as mean squares (MS). An approximation of s, very useful for small samples, is the range: R = max-min observed value.

The estimate of the covariance is given by:

$$s_{xy} = \frac{1}{n-1} \sum(x_j - \bar{x})\,(y_j - \bar{y})$$

and of the linear correlation coefficient, by:

s.

The sampling variance and the correlation coefficient need the knowledge of higher order moments of the population distribution. Simple formulas are available only if the population is normal or at least approximately normal (see also Section 2.4).

The variance of s^2, with normal parent distribution, is: $\frac{2\sigma^4}{n}$; of s it is: $\frac{\sigma^2}{2n}$; and of r: $\frac{(1-\rho^2)^2}{n}$.

1.4 Composite Sample Statistics

The scope of sampling is the estimation of the population parameters with the greatest accuracy possible. It is possible to increase the efficiency of estimations by a convenient sampling plan that evaluates independently different factors at different levels which may influence observations about individuals. This influence is called the effect of the factor level. Levels are the different values which each factor can assume, e.g. if one factor is time, the different dates are the levels. As a consequence, besides a smaller sampling variance, information about the importance of the factors in terms of variability, measured by their contribution to the total variance, can be obtained. The contributions to the total variance are estimated by analyses of variance (ANOVA).

The levels of the factors can be fixed or random. In the first case the main interest lies in estimating and testing hypotheses about (i.e. inference) the effects of the factor levels in the model. These effects are what we call fixed effects. In the second case the levels are thought to be a random sample out of a theoretical population of levels and the prime interest lies in estimating the variances of this population. Effects of this nature are called random effects and the variances associated with random effects, variance components. In considering the difference between fixed and random models, the important question is that of inference: when inferences are going to be confined to the effects in the model, the effects are considered fixed; when inferences will be made about a population of effects from which those in the data are considered to be a random sample, then the effects are considered as random.

For any balanced random model the method of estimating variance components relies on the MS of the ANOVA for the fixed effects model.

For unbalanced data there are several methods each of which simplifies to the method used for balanced data.

1.4.1 One factor models

One way of obtaining composite samples is that to consider one factor with p levels $(\theta_1..\theta_i..\theta_p)$. For each level a sample of measurements on n_i $(i = 1..p)$ individuals is obtained. The result of such a sampling design can be summarized in a two way table, where also the fundamental sample functions are reported (Table II).

Table II

A one factor sampling plan

Samples		Factor levels θ_i $i = 1 \ldots i = p$	Overall
Observations	$j = 1$	$y_{11} \cdots y_{p1}$	
	$j = n_i$	$y_{1n_1} \cdots y_{pn_p}$	
Number of individuals per sample n_i		$n_1 \cdots n_p$	$n = \sum n_i$
Sample mean $\bar{y}_i$		$\bar{y}_1 \cdots \bar{y}_p$	$\bar{y} = \frac{1}{n} \sum \bar{y}_i n_i$
Sample SS		$SS(\varepsilon_1) \cdots SS(\varepsilon_p)$	$SS(\varepsilon) = \sum SS(\varepsilon_i)$
Sample variance s_i^2		$s_1^2 \cdots s_p^2$	$s_\varepsilon^2 = \frac{1}{n-p} SS(\varepsilon)$

The model of such a design can be written as follows:

$$y_{ij} \quad \mu + \theta_i + \varepsilon_{ij} \quad \begin{cases} i = 1 \ldots p \\ j = 1 \ldots n_i \end{cases}$$

where ε_{ik} indicates the error or the residual variation term. That means if the factor had not any influence on the observations then $\theta_i = 0$ for all i and the expected value of y_{ij} would be the overall mean μ. If it had, then for the i^{th} level the expected value of y_{ij} would be $\mu + \theta_i$. The difference between the observation y_{ij} and the expected value $\mu + \theta_i$ is the error or residual term, i.e. $y_{ij} - (\mu + \theta_i) = \varepsilon_{ij}$.

If homoscedasticity (i.e. if all variances s_i^2 do not differ significantly between each other: tests of Bartlett, Cochran, Scheffé*) can be assumed, a pooled estimate of the population variance between individuals within each sample can be obtained by the MS:

$$s_\varepsilon^2 = \frac{\Sigma (n_i - 1) s_i^2}{\Sigma n_i - p} .$$

* See Chapter 2.4.3

This MS is often called the residual or the within (samples) or the error MS, and it is an estimate of the population error variance σ_ε^2.

On the other hand, considering only the p sample means $\bar{y}_i$, we then obtain the SS due to the factor after eliminating the influence due to the mean μ:

$$SS\ (\theta_i \mid \mu) = \sum (\bar{y}_i - \bar{y})^2 n_i$$

$$MS\ (\theta_i \mid \mu) = \frac{\sum (\bar{y}_i - \bar{y})^2 n_i}{p-1}$$

This MS is thus a measure of the influence of the factor, and it is an estimate of:

$$\sigma_\varepsilon^2 + \frac{1}{p-1} \sum n_i \theta_i^2$$

in the fixed effects model, or

$$\sigma_\varepsilon^2 + n_I \sigma_I^2$$

in the random effects model, where

$$n_I \quad \frac{1}{p-1} \left(n - \frac{\sum n_i^2}{n} \right)$$

or $n_I = \frac{n}{p} = m$, if all n_i are equal, $n_i = m$. σ_I^2 indicates the variance associated with the factor.

Usually the results of such an analysis are summarized in a table, called analysis of variance (ANOVA) of the model $\mu + \theta_i + \varepsilon_{ij}$, like Table III, where E[MS] stands for expected MS. It is now possible to compare $MS(\theta_i \mid \mu)$ with the residual MS by the F-test, with p-1 and n-p degrees of freedom (DF):

$$F_{p-1,n-p} - \frac{MS\ (\theta_i \mid \mu)}{s_\varepsilon^2}$$

which enables us to decide, at a given level of significance, if the effect of the factor affects the sample means.

Now we have two possibilities:

(i) The means are not significantly different
(ii) The means are significantly different (*).

(*) If the means are significantly different, it may be that in one sample there are some observatior which differ greatly from the others. The problem of deciding whether one or more of a set of observations has come from a different population from that generating the other observations, leading thus to different means faces every practical statistician, and perhaps, indeed, most practical scientists. It is the problem of outlying observations or outliers (see Chapter 2.8.1). Sometimes it is possible to replace the value with a corrected value or to repeat the observation. If no reason for the unusual value can be found, the problem is more difficult and no unique solution can be suggested, but the one to eliminate observations in the statistical analysis and to report them along with the statistical analysis. This allows the reader of the report to judge for himself; if he wishes he can replace the outliers and rerun the analysis.

Table III

ANOVA of model $\mu + \theta_i + \varepsilon_{ij}$

	Source of variation	SS	DF	MS	E[MS] Expected MS
Model	Mean	$SS(\mu) = n\bar{y}^2$	1		
	Factor (I after μ)	$SS(\theta_i \mid \mu)$	p-1	$MS(\theta_i \mid \mu)$	$\sigma_\varepsilon^2 + n_I \sigma_I^2$
	Residual	$SS(\varepsilon)$	n-p	$MS(\varepsilon)$	σ_ε^2
	TOTAL:	SS	n		

Under the first circumstance, the population mean can be estimated directly by $\bar{y}$ with a sampling (standard) error of:

$$s_{\bar{y}} = \frac{s_\varepsilon}{\sqrt{n}} .$$

Under the second circumstance, the estimation of a unique mean is arbitrary if the frequency N_i of the individuals of the population subject to the single levels of the factor are not known. If they are, or can be guessed through the sample, then the estimation of the mean is obtained by:

$$\bar{y}' = \frac{1}{N} \sum N_i \bar{y}_i \quad \text{with} \quad N = \sum N_i .$$

If the sample frequencies can be supported to be proportional to the analogous in the population, then $\bar{y}'$ reduces to $\bar{y}$.

The standard error of $\bar{y}$ for $N_i >> n_i$ is then:

$$\sigma_{\bar{y}} = \sqrt{\frac{SS\ (\theta_i \quad \mu)}{n}} \qquad (1.4\text{-}1)$$

In the case of random effects models, if the means are not significantly different σ_I^2 may be put equal to zero; otherwise it can be estimated solving the system:

$$\begin{cases} MS(\theta_i \mid \mu) = \sigma_\varepsilon^2 + n_I \sigma_I^2 \\ MS(\varepsilon) = \sigma_\varepsilon^2 \end{cases}$$

Hence,

$$\hat{\sigma}_I^2 = S_I^2 = \frac{MS(\theta_i \mid \mu) - MS(\varepsilon)}{n_I}$$

1.4.2 Two factor nested model

The extension of the one factor model includes one or more other factors. We will briefly discuss only the situation of one more factor, since all other and more complex designs are based on these. We restrict ourselves to the situation of a two-way nested design. This is obtained when we consider one factor with p levels as for the single factor design, and for each level a second factor at q_i levels.

The n_i observations of the sample for the i^{th} level of the first factor are classified in q_i sub-samples, with frequencies n_{ij}, $j = 1 \ldots q_i$, $\sum_j n_{ij} = n_i$. In accordance with the notational convention the original frequencies n_i will now be replaced by $n_{i.}$.

Such a classification is called a two-way nested or hierarchical classification, which means that the classification of the individuals is nested within the samples obtained for the different factor levels.

The model is now:

$$y_{ijk} = \mu + \theta_i + \theta_{j(i)} + \varepsilon_{ijk} \qquad \begin{cases} i = 1 \ldots p \\ j = 1 \ldots q_i \\ k = 1 \ldots n_{ij} \end{cases} \qquad (1.4\text{-}2)$$

where θ_i specifies the effect of the first factor and $\theta_{j(i)}$ that one of the second nested within the first.

The estimate of the mean is readily obtained if the population frequencies are known or estimable through the sample frequencies:

$$\frac{1}{N} \sum\sum N_{ij}\, y_{ij} \quad \text{or} \quad \frac{1}{n} \sum\sum n_{ij}\, y_{ij} \,.$$

The MS and expected MS due to the factors and the residual are obtainable as:

$$
\left.
\begin{aligned}
MS(\theta_i) &= \frac{1}{p-1} \sum (\bar{y}_{i\cdot} - \bar{y})^2 n_i \; ; \qquad E[MS(\theta_i)] = \sigma_\varepsilon^2 + n_{II}\sigma_{II}^2 + n_I\sigma_I^2 \\
MS(\theta_{j(i)}) &= \frac{1}{\Sigma q_i - p} \sum\sum (\bar{y}_{ij} - \bar{y}_{i\cdot})^2 n_{ij} \; ; \quad E[MS(\theta_{j(i)})] = \sigma_\varepsilon^2 + n'_{II}\,\sigma_{II}^2 \\
\sigma_\varepsilon^2 &= \frac{1}{n - \Sigma q_i} \sum\sum\sum (y_{ijk} - \bar{y}_{ij})^2
\end{aligned}
\right\} \quad (1.4\text{-}3)
$$

with,

$$
\left.
\begin{aligned}
n_{II} &= \frac{1}{p-1} \left[\sum_i^p \frac{\sum_j^{q_i} n_{ij}^2}{n_{i\cdot}} - \frac{\sum\sum n_{ij}^2}{n_{\cdot\cdot}} \right] \\
n'_{II} &= \frac{1}{\Sigma q_i - p} \left[n_{\cdot\cdot} - \sum_i^p \frac{\sum_j^{q_i} n_{ij}^2}{n_{i\cdot}} \right] \\
n_I &= \frac{1}{p-1} \left[n_{\cdot\cdot} - \frac{\sum_i n_{i\cdot}^2}{n_{\cdot\cdot}} \right] .
\end{aligned}
\right\} \quad (1.4\text{-}4)
$$

If n_{ij} = constant = m, and q_i = constant = q, then we have $n_{II} = n'_{II} = m$ and $n_I = mq$.

Hypotheses about the significance of the two factors can be tested by the ratios:

$$\frac{MS(\theta_i)}{MS(\theta_{j(i)})} \quad \text{and} \quad \frac{MS(\theta_{j(i)})}{\hat{\sigma}_\varepsilon^2} \; .$$

1.5 Regression or covariance models

Another way to obtain more information out of a set of observations, is to consider the dependence of the variable studied, y, upon one or more other variables, x,w, ... which are concomitant to the first. This leads to the regression analysis and the analysis of covariance which will only be briefly outlined.

The general model is that of an ANOVA which will be broadened to:

$$y_{ij} = \mu + \theta_i + \beta x_{ij} + \varepsilon_{ij} \qquad (1.5\text{-}1)$$

where x_{ij} are the values of the concomitant variable, and β is a measure of the dependence of y on x (regression coefficient). This model can be thought of as a synthesis of two models, viz. the one-way classification:

$$y_{ij} = \mu + \theta_i + \varepsilon_{ij} \qquad (1.5\text{-}2)$$

and the simple regression model:

$$y_{ij} = \mu + \beta x_{ij} + \varepsilon_{ij} \qquad (1.5\text{-}3)$$

for which the ANOVA of Table IV can be considered (Searle, 1971).

The left half of the table leads to the test of the hypothesis whether, after having estimated μ and θ_i, the regression on x indicated by $\beta|\mu,\theta_i$ introduced in the first model (1.5-1) is significant or not.

The right half of the table leads to the test of the hypothesis whether the factor effects θ_i are significant after having removed the influence of μ and $\beta|\mu$.

Generalizations of model (1.5-1) consider different regression coefficients β_i instead of only one for all factor levels. This is the so-called <u>intraclass regression</u> model. In that case $SS(\beta|\mu,\theta_i)$, with 1 DF must be substituted by:

$$SS\,(\beta_i|\mu,\theta_i) = \sum_i \frac{[\sum (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i)]^2}{\sum (x_{ij} - \bar{x}_i)^2}, \text{ with p, DF}$$

and $SS\,(\beta|\mu)$ by:

$$SS\,(\beta_i|\mu) = \sum_i \frac{[\sum (x_{ij} - \bar{x})(y_{ij} - \bar{y})]^2}{\sum (x_{ij} - \bar{x})^2}, \text{ with (p-1), DF.}$$

Hence, the DF for the residual SS are reduced to $n - 2p$.

Table IV

ANOVA for the regression model - degrees of freedom in brackets [DF]

$\beta \mid \mu, \theta_i$	MODELS	$\theta_i \mid \mu, \beta$
$SS(\mu) = n\bar{y}^2$ [1]		
$SS(\theta_i \mid \mu) = \sum(\bar{y}_i - \bar{y})^2 n_i$ [p-1]		$SS(\beta \mid \mu) = \frac{\{\sum\sum (y_{ij} - \bar{y})(x_{ij} - \bar{x})\}^2}{\sum\sum (x_{ij} - \bar{x})^2}$ [1]
$SS(\mu, \theta_i) = \sum \bar{y}_i^2 n_i$ [p]		$SS(\mu, \beta) = SS(\mu) + SS(\beta \mid \mu)$ [2]
$SS(\beta \mid \mu, \theta_i) = \frac{\{\sum\sum (y_{ij} - \bar{y}_i)(x_{ij} - \bar{x}_i)\}^2}{\sum\sum (x_{ij} - \bar{x}_i)^2}$ [1]		$SS(\theta_i \mid \mu, \beta) = SS(\theta_i \mid \mu) + SS(\beta \mid \mu, \theta_i) - SS(\beta \mid \mu)$ [p-1]
Residual	$SS = \sum\sum (y_{ij} - \bar{y}_i)^2 - SS(\beta \mid \mu, \theta_i)$	[n-p-1]
TOTAL:	$SS = \sum\sum y_{ij}^2$	[n]

1.6 The problem of the sample size

1.6.1 One sample

In determining the sample size one has to make statements about:

(a) The precision ε of the estimates

(b) The confidence level $1-\alpha$, with which decisions are taken.

Hence, if $\phi(x)$ is the true and unknown population parameter, $f(x)$ its estimate obtained through a sample, above statements can be summarized as:

$$P\left[\,|\, f(x) - \phi(x) \,|\, < \varepsilon\right] = 1 - \alpha\,. \qquad (1.6\text{-}1)$$

This means that the probability to obtain distances of deviations between the estimate and the true value smaller than ε, is $1-\alpha$. In general to solve (1.6-1) the frequency distribution of the population must be known or assumed. If the estimate is the mean $\bar{x}$ or a proportion $\frac{k}{n}$ of the sample, then (1.6-1) can be easily solved, in most cases, even without knowing the exact distribution. In case of the mean, we have Student's test (2.3-2):

$$t_{n-1} = \frac{\bar{x} - \mu}{\sqrt{\frac{N-n}{N} \frac{s^2}{n}}} \qquad (1.6\text{-}2)$$

where the factor $\frac{N-n}{N}$ is the finite population correction of the variance and can be omitted if N is much larger than the expected sample size. In (1.6-1) clearly $\varepsilon = \bar{x}-\mu$, so that on taking the critical t value at the confidence level of $1-\frac{\alpha}{2}$, i.e. $t_{(1-\frac{\alpha}{2})}$, we have:

$$\varepsilon = t_{(1-\frac{\alpha}{2})} \sqrt{\frac{N-n}{N} \frac{s^2}{n}} \qquad (1.6\text{-}3)$$

for a finite population, or

$$\varepsilon = t_{(1-\frac{\alpha}{2})} \frac{s}{\sqrt{n}} \qquad (1.6\text{-}4)$$

for large populations or universe. Hence we have:

$$n = \frac{\left(\frac{t_{(1-\frac{\alpha}{2})} s}{\varepsilon}\right)^2}{1 + \frac{1}{N}\left(\frac{t_{(1-\frac{\alpha}{2})} s}{\varepsilon}\right)^2} \qquad (1.6\text{-}5)$$

for finite populations, or

$$n = \left(\frac{t_{(1-\frac{\alpha}{2})} s}{\varepsilon}\right)^2 \qquad (1.6\text{-}6)$$

for large populations or universe.

If neither $t_{(1-\frac{\alpha}{2})}$ nor s are known in advance they must be guessed. This is best done by guessing the ratio $w = \varepsilon / t_{(1-\frac{\alpha}{2})}$ which is the desired standard error of the estimate $\bar{x}$, obtaining $n = (\frac{s}{w})^2$. The sample variance s^2 can only be estimated from previous sampling of a similar or related population, or by intelligent guess-work (Cochran, 1960). In practice informations about s^2 are obtained by a pilot sampling. Then $t_{(1-\frac{\alpha}{2})}$ can be substituted by the normal approximation $t_{n(1-\frac{\alpha}{2})} = z_{(1-\frac{\alpha}{2})}\sqrt{\frac{n}{n-2}}$, or much easier by the critical value of t for the smallest sample size expected, or iteratively.

1.6.2 Two samples

We consider now the case of determining the sizes of two samples for which the means should be compared by the test (2.3-9):

$$t_{n_1+n_2-2} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \left(\frac{n_1+n_2}{n_1 n_2}\right)}} \qquad (1.6\text{-}7)$$

We analyse two situations:

(i) n_1 fixed, n_2 variable

(ii) both n_1, n_2 variable, but equal, $n_1 = n_2 = m$.

In the first case, putting $\bar{x}_1 - \bar{x}_2 = \varepsilon$ and solving (1.6-7) for n_2 we obtain:

$$n_2 \geq \frac{1}{2A}\left(B + \sqrt{B^2 + 4AC}\right) \qquad (1.6\text{-}8)$$

where,

$$A = \varepsilon^2 n_1 - t^2 s_2^2 \text{, which must be } > 0$$

$$B = t^2(n_1 - 1)(s_1^2 + s_2^2) - \varepsilon^2 n_1(n_1 - 2)$$

$$C = t^2 n_1[s_1^2(n_1 - 1) - s_2^2] \text{ .}$$

In the second case,

$$m = \frac{t^2(s_1^2 + s_2^2)}{\varepsilon^2} \text{ .} \qquad (1.6\text{-}9)$$

1.7 Examples

In the monitoring pilot projects, four different sources of variation (i.e. factors) can be distinguished (Bernhard, 1976):

1. Variability of the chemical analytical method (between determinations)
2. Variability of the chemical pretreatment (between pretreatments)
3. Variability caused by intrinsic (genetic) biological factors (between "samples", same site and same date)
4. Variability caused by environmental factors (between sites and date or between dates, same sites).

Now suppose we consider nine samples of *Mullus barbatus*, which differ from environmental factors, i.e. they differ with sites and time - factor 4 (Stöppler et al.,i.pr.)

Each sample is composed of a number of individuals (fishes) which differ from factor 3.

For all individuals of the first sample (see Table V) two pretreatments have been performed with two determinations.

For all other samples (2 to 9) data for only one pretreatment with two determinations are given. There is only one fish in sample No.2 with a missing observation.

For all samples, except No.4, lengths and weights are given for each fish.

The resulting observations do not fit in a balanced scheme, since not all samples have data for all factor levels.

Although nowadays the high-speed computer can analyse these models, though with some approximations, we suggest for the sake of clearness, to follow a stepwise analysis, treating at the beginning all samples separately.

Table V

Concentrations y_{ijhk} of a contaminant, lengths and weights for 9 samples of *Mullus barbatus*

Sample Number	Individual Number	y_{ijhk}				x_{ij} length	w_{ij} weight
		Pretreatment h=1		Pretreatment h=2			
		Determination		Determination			
		k=1	k=2	k=1	k=2		
1	01	0.037	0.039	0.046	0.044	15.3	75.19
	02	0.266	0.264	0.270	0.267	13.3	47.61
	03	0.135	0.139	0.159	0.159	10.8	23.77
	04	0.095	0.100	0.112	0.109	11.5	27.44
	05	0.101	0.102	0.116	0.114	11.3	28.24
	06	0.213	0.213	0.215	0.215	12.3	34.24
	07	0.228	0.203	0.219	0.219	11.9	29.05
	08	0.167	0.185	0.186	0.186	13.6	50.01
	09	0.766	0.750	0.686	0.678	12.8	37.01
	10	0.054	0.054	0.046	0.040	14.2	55.66
2	01	0.252	0.246			16.0	73.58
	02	0.264	0.258			13.0	40.99
	03	0.149	0.146			14.3	52.03
	04	0.308				13.6	45.91
	05	0.271	0.247			12.8	34.43
	06	0.315	0.305			12.5	35.04
	07	0.163	0.152			13.5	42.57
	08	0.380	0.379			14.2	51.96
	09	0.109	0.103			13.0	41.72
	10	0.072	0.072			12.3	36.53
3	01	0.228	0.224			14.0	63.3
	02	0.243	0.228			14.5	59.6
	03	0.209	0.211			14.2	57.9
	04	0.184	0.184			13.2	46.8
	05	0.168	0.164			14.3	57.9
4	01	0.066	0.062				
	02	0.081	0.081				
	03	0.073	0.071				
	04	0.071	0.074				
	05	0.058	0.059				
	06	0.048	0.049				
	07	0.045	0.045				
	08	0.088	0.086				
	09	0.118	0.125				
	10	0.045	0.041				
5	01	0.062	0.062			16.5	67.0
	02	0.051	0.057			14.0	46.23
	03	0.043	0.044			15.0	53.5
	04	0.081	0.082			15.0	56.1
	05	0.118	0.122			15.0	59.59

Continued

Sample Number	Individual Number	y_{ijhk} Pretreatment h=1 Determination k=1	y_{ijhk} Pretreatment h=1 Determination k=2	y_{ijhk} Pretreatment h=2 Determination k=1	y_{ijhk} Pretreatment h=2 Determination k=2	x_{ij} length	w_{ij} weight
6	01	0.051	0.051			13.0	31.93
	02	0.096	0.097			13.0	32.15
	03	0.144	0.167			14.5	43.39
	04	0.218	0.202			14.5	55.16
	05	0.102	0.104			14.5	38.61
7	01	0.331	0.331			11.0	19.49
	02	0.280	0.287			10.5	17.34
	03	0.173	0.180			12.5	27.07
	04	0.135	0.132			12.0	25.59
	05	0.115	0.127			13.5	32.77
8	01	0.420	0.410			16.0	73.40
	02	0.557	0.555			16.0	72.00
	03	0.115	0.110			17.0	92.42
	04	0.179	0.186			11.0	26.67
	05	0.168	0.167			11.0	24.29
	06	0.154	0.154			10.5	19.18
	07	0.089	0.102			11.0	18.71
	08	0.269	0.166			14.0	50.57
	09	0.258	0.262			11.5	25.64
	10	0.257	0.242			16.0	85.47
9	01	0.085	0.083			10.0	20.00
	02	0.147	0.147			8.5	10.90
	03	0.092	0.092			10.0	17.26
	04	0.189	0.192			9.0	12.59
	05	0.050	0.052			10.5	22.05
	06	0.030	0.030			8.0	9.94
	07	0.027	0.021			10.5	25.74
	08	0.047	0.054			10.0	21.29
	09	0.024	0.024			9.5	17.52
	10	0.116	0.123			9.5	16.61
	11	0.046	0.050			9.0	12.01
	12	0.053	0.054			9.0	15.91
	13	0.086	0.086			9.0	15.91
	14	0.060	0.064			10.5	19.24
	15	0.077	0.079			12.5	35.21
	16	0.049	0.050			10.0	18.06
	17	0.164	0.173			9.5	13.46
	18	0.073	0.078			9.5	14.64
	19	0.087	0.091			12.0	29.27
	20	0.054	0.055			9.5	15.05
	21	0.092	0.096			10.5	23.00

1.7.1 Example of a one-factor model

We consider sample No.2 (Table V) of p = 10 individuals (= specimen), ignoring lengths and weights, and set up Table VI following Table II determining all parameters. Since for nine individuals we have only two observations (= determinations), evidently the sample SS is equal to the sample variance. In the case of the fourth level (= specimen) we have only one determination so that the sample variance cannot be evaluated. Therefore, this specimen does not bring any contribution to the error or residual MS, which in this case is equal to the determination MS.

Homoscedasticity is tested by Cochran's test (2.4-12):

$$g_{9,2} = \frac{s^2[9]}{\sum s_i^2} = \frac{28.8}{45.75} = 0.63$$

At the 5% significance level, we have from Figure 6* $g_{9,2(.95)} = .64$ which is larger than the observed value; hence, with some optimism, the variances can be considered as homogeneous. This is due especially to the fact that all variances of Table VI have only 1 DF.

Table VI

A one factor (= individuals) sampling plan

y_{ij}	Factor levels i										Overall
	1	2	3	4	5	6	7	8	9	10	
j = 1	0.252	0.264	0.149	0.308	0.271	0.315	0.163	0.380	0.109	0.072	
j = 2	0.246	0.258	0.146	-	0.247	0.305	0.152	0.379	0.103	0.072	
n_i	2	2	2	1	2	2	2	2	2	2	19
$\bar{y}_i$	0.249	0.261	0.1475	0.308	0.259	0.310	0.1575	0.3795	0.106	0.072	0.22058
Sample SS($\times 10^{-5}$)	1.8	1.8	0.45	0	28.8	5	6.05	0.05	1.8	0	45.75
$s_i^2(\times 10^{-5})$	1.8	1.8	0.45	-	28.8	5	6.05	0.05	1.8	0	$s_\epsilon^2 = 5.08339$

* See page 57

Further, we obtain SS due to the individuals,

$$SS\ (\theta_i|\mu) = 0.171029 \quad \text{with} \quad MS\ (\theta_i|\mu) = 0.019$$

So that,

$$F = \frac{MS\ (\theta_i|\mu)}{s_\varepsilon^2} = 373.8$$

which is extremely high, stating that the variability between individuals is incomparably higher than that of determinations.

If the mean $\bar{y}$ can be accepted as an estimate of the population mean μ, the sampling error must be calculated by (1.4-1):

$$s_{\bar{y}} = \sqrt{\frac{0.171029}{19}} = 0.095 \ .$$

1.7.2 Example of a regression test

Now we also consider the lengths of individuals of sample No.2 (Table V) and evaluate the regression of the individual measures of Table VI on the lengths.

We immediately have:

$$s_{xy} = \frac{1}{n-1} \sum\sum (x_i - \bar{x})(y_{ij} - \bar{y}) = \frac{1}{n-1} \sum (x_i - \bar{x})(\bar{y}_i - \bar{y})\, n_i$$

$$= \frac{1}{n-1} \left[\sum x_i \bar{y}_i n_i - n \bar{x} \bar{y} \right]$$

$$= \frac{1}{18} \left[57.099 - 19 \times 13.5 \times 0.22058 \right] = 0.0266$$

$$s_x^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 n_i = \frac{1}{n-1} \left[\sum x_i^2 n_i - n\bar{x}^2 \right]$$

$$= \frac{1}{18} \left[3492.48 - 3467.88 \right] = 1.367$$

And analogously,

$$s_y^2 = \frac{1}{18} \left[1.095933 - 0.92445 \right] = 0.00953$$

Thus, the correlation coefficient between the concentration measures and the length in the example considered is $r = .232$. To test if this value is sufficiently high to assure a dependence of the concentration measures y_{ij} on the length x_i, we must apply test (2.5-2) obtaining

$$t_{17} = \frac{.232}{\sqrt{1 - (.232)^2}} \sqrt{17} = .963$$

This lies well below the critical value $t_{17(.975)} = 2.11$ (Table A-4) so that we may conclude that there is no evidence that a dependence between y_{ij} and x_i does exist, and that it is hence meaningless to evaluate a regression line.

1.7.3 Example of a two-factor nested model with the determination of the variance components for individuals and pretreatment

We consider only sample No.1 (Table V), ignoring lengths and weights. With the remaining data we can set up a 3-way nested table: determinations nested within pretreatments, which are nested within individuals. The model can therefore be written as follows:

$$y_{1jhk} = \mu_1 + \theta_{1j} + \theta_{h(1j)} + \theta_{k(1jh)}$$

or dropping the sub-script 1,

$$y_{jhk} = \mu + \theta_j + \theta_{h(j)} + \theta_{k(jh)}$$

$$\mu + \theta_j + \theta_{h(j)} + \varepsilon_{jhk}$$

which is like model (1.4-2).

Since there are no replications, the last term $\theta_{k(jh)}$ is confounded with the error term ε_{jhk}. Consequently, the model can be treated as a 2-way nested model with replications. The complete ANOVA is reported in Table VII. In the last column are the MS due to the factors which all differ significantly.

Table VII

ANOVA of model $y_{jhk} = \mu + \theta_j + \theta_{h(j)} + \theta_{k(jh)}$

Source of variation		SS	DF	MS
Individuals	θ_j	SS (θ_j) = 1.385894	9	0.153988
Pretreatments	$\theta_{h(j)}$	SS $(\theta_{h(j)})$ = 0.007689	10	0.000769
Determinations	$\theta_{k(jh)}$	SS $(\theta_{k(jh)})$ = 0.000532	20	0.000027
Deviance		sum 1.394115	39	
Mean	μ = 0.205225	SS (μ) = 1.684692	1	
TOTAL:		sum 3.078807	40	

It is thus possible to evaluate the variance components from (1.4-3):

$$\left.\begin{aligned} 0.000027 &= \sigma^2_\varepsilon \\ 0.000769 &= \sigma^2_\varepsilon + 2\sigma^2_{II} \\ 0.153988 &= \sigma^2_\varepsilon + 2\sigma^2_{II} + 4\sigma^2_I \end{aligned}\right\} \rightarrow \left\{\begin{aligned} s^2_\varepsilon &= 0.000027 \\ s^2_{II} &= 0.000374 \\ s^2_I &= 0.038305 \end{aligned}\right.$$

Hence it results that the variance due to individuals is more than 100 times larger than that due to pre-treatments, which in turn is almost 14 times larger than that due to determinations.

1.7.4 Example of a two-factor nested model with the determination of the variance components for samples and individuals

Now we consider sample Nos. 2 to 9 (Table V), still ignoring lengths and weights, excluding fish i=2, j=4 because of the missing observation. The aim of this analysis is to estimate the MS between samples, individuals and determinations. The model is now without sub-script h,

$$y_{ijk} = \mu + \theta_i + \theta_{j(i)} + \theta_{k(ij)}$$
$$\mu + \theta_i + \theta_{j(i)} + \varepsilon_{ijk}$$

with ANOVA reported in Table VIII.

Table VIII

ANOVA of model $y_{ijk} = \mu + \theta_i + \theta_{j(i)} + \theta_{k(ij)}$

Source of variation		SS		DF	MS
Samples	θ_i	SS (θ_i)	= 0.698404	7	0.099772
Individuals	$\theta_{j(i)}$	SS ($\theta_{j(i)}$)	= 0.737255	62	0.011891
Determinations	$\theta_{k(ij)}$	SS ($\theta_{k(ij)}$)	= 0.006948	70	0.000099
Deviance		sum	1.442607	139	
Mean	μ = 0.139421	SS (μ)	= 2.721366	1	
TOTAL:		sum	4.163973	140	

Again the MS differ significantly, so that we can evaluate the variance components by (1.4-3) obtaining:

$$n_{II} = n'_{II} = m = 2$$

$$n_I = \frac{1}{p-1}\left[n\cdot\cdot - \frac{m^2 \sum q_i^2}{n\cdot\cdot}\right] = \frac{1}{7}\left[140 - \frac{3\,288}{140}\right] = 16.64$$

Hence,

$$\left.\begin{aligned} 0.000099 &= \sigma_\varepsilon^2 \\ 0.011891 &= \sigma_\varepsilon^2 + n'_{II}\,\sigma_{II}^2 \\ 0.099772 &= \sigma_\varepsilon^2 + n_{II}\,\sigma_{II}^2 + n_I\,\sigma_I^2 \end{aligned}\right\} \rightarrow \left\{\begin{aligned} s_\varepsilon^2 &= 0.000099 \\ s_{II}^2 &= 0.005896 \\ s_I^2 &= 0.005281 \end{aligned}\right.$$

It is seen that the variances due to samples and individuals are almost equal, but they are considerably (59 times) larger than that due to determinations.

Comparing these results with those of Table VII, significant differences must be observed, both in determination and individual MSs and variance components. This is probably due to some outliers and/or some wrong determinations. Hence a synthesis of both tables is not advisable.

1.7.5 Example of a covariance model

Now we eliminate sample No.4 (Table V) because there are no measurements about lengths and weights of individuals, and we determine the mean concentration $\bar{y}_{ij}$ for all individuals, so that we now have 3 values: $\bar{y}_{ij}$, x_{ij} and w_{ij}.

The models which can be analysed separately are:

$$\bar{y}_{ij} = \mu + \bar{\theta}_i + \beta x_{ij} + \varepsilon_{ij}$$

and,

$$\bar{y}_{ij} = \mu + \bar{\theta}_i + \gamma w_{ij} + \varepsilon_{ij}$$

The corresponding ANOVAs are reported in Tables IX and X. It can be seen that neither the introduction of x_{ij} nor that of w_{ij} are of significant importance, since the residual MS remain almost unaltered, reducing only from 0.01139 without the concomitant variables to 0.01134 considering x_{ij} and even increasing to 0.01143 considering w_{ij}.

We may thus conclude that the concentration does not depend significantly on the length and weight of the fishes. But it must be underlined that in the samples considered here, lengths and weights did not differ very much so that one could hardly expect another result. Regression models are only meaningful if the regressor, i.e. the independent variable has a variance as large as possible.

Instead of the above models, it is possible to include x_{ij} and w_{ij} contemporarily to obtain the model,

$$\bar{y}_{ij} = \mu + \bar{\theta}_i + \beta x_{ij} + \gamma w_{ij} + \varepsilon_{ij} .$$

Also, in this case, the introduction of the concomitant variables does not reduce the error MS which instead increases to 0.01147, so that the ANOVA table is not reported here.

Table IX

ANOVA of the regression model $\bar{y}_{ij} = \mu + \bar{\theta}_i + \beta x_{ij} + \varepsilon_{ij}$ [DF]

$\beta \mid \mu, \bar{\theta}_i$	MODELS	$\bar{\theta}_i \mid \mu, \beta$
SS (μ) = 1.839023 [1]		
SS ($\bar{\theta}_i \mid \mu$) = 0.331211 [7]		SS ($\beta \mid \mu$) = 0.108838 [1]
SS ($\mu, \bar{\theta}_i$) = 2.170234 [8]		SS (μ, β) = 1.947861 [2]
SS ($\beta \mid \mu, \bar{\theta}_i$) = 0.014385 [1]		SS ($\bar{\theta}_i \mid \mu, \beta$) = 0.236758 [7]
Residual SS = 0.703327 [62]		
TOTAL: SS = 2.887946 [71]		

Table X

ANOVA of the regression model $y_{ij} = \mu + \bar{\theta}_i + \gamma w_{ij} + \varepsilon_{ij}$ [DF]

$\gamma \mid \mu, \bar{\theta}_i$	MODELS	$\bar{\theta}_i \mid \mu, \gamma$
SS (μ) = 1.839023 [1]		
SS ($\bar{\theta}_i \mid \mu$) = 0.331211 [7]		SS ($\gamma \mid \mu$) = 0.096325 [1]
SS ($\mu, \bar{\theta}_i$) = 2.170234 [8]		SS (μ, γ) = 1.935342 [2]
SS ($\gamma \mid \mu, \bar{\theta}_i$) = 0.009054 [1]		SS ($\bar{\theta}_i \mid \mu, \gamma$) = 0.243940 [7]
Residual SS = 0.708658 [62]		
TOTAL: SS = 2.887946 [71]		

1.7.6 Example of a comparison of three regression lines

We analyse the data y_{ij} obtained by Renzoni (1977) for the years 1973 (i=1), 1975 (i=2) and 1976 (i=3) about the mercury content in small-scaled Scorpion fishes *(Scorpaena porcus)* with one concomitant variable, w_{ij} (body weight), Table XI.

Table XI

Scorpaena porcus - Mercury content (white muscle) y_{ij} and body weight w_{ij}

j	1973		1975		1976	
	y_{1j}	w_{1j}	y_{2j}	w_{2j}	y_{3j}	w_{3j}
1	1.61	28.	1.6	27.2	1.23	78.6
2	1.93	28.	1.24	28.2	3.046	85.
3	1.52	36.	1.25	28.4	1.468	79.
4	1.64	37.	1.1	29.9	2.514	87.5
5	1.55	37.	1.07	29.9	2.518	82.4
6	1.64	38.	1.41	30.1	1.776	83.6
7	1.72	38.	1.1	30.7	2.084	78.6
8	1.3	38.	1.12	31.2	2.067	79.
9	2.28	38.	0.93	31.3	1.328	82.
10	2.16	38.	1.02	31.7	1.433	82.5
11	1.68	39.	1.05	32.8	1.343	33.1
12	1.65	40.	1.13	36.5	1.213	35.1
13	1.7	40.	0.96	37.1	1.004	35.2
14	1.7	41.	1.20	37.4	1.852	55.5
15	1.9	41.	1.26	38.5	1.855	57.5
16	1.85	43.	1.11	39.1	1.555	72.7
17	2.12	45.	1.15	43.	1.245	62.
18	1.64	47.	1.41	45.2	2.297	63.
19	1.9	47.	1.27	45.5	2.233	65.
20	1.92	48.	1.09	45.7	1.457	78.
21	2.0	54.	1.5	46.8	1.454	67.7
22	2.14	50.	1.85	52.4	1.571	66.
23	2.66	60.	1.45	54.1	3.805	74.1
24	1.89	60.	1.48	56.7	1.485	63.2
25	2.22	63.	1.41	57.5	1.176	87.2
26	2.0	64.	1.04	60.1	1.729	74.9
27	2.04	66.	1.35	60.6	1.795	80.2
28	2.45	66.	1.31	65.	.893	92.2
29	2.0	69.	1.26	65.5	1.663	87.9
30	2.18	70.	0.85	68.1	1.805	91.1
31	2.3	72.	1.18	70.7	1.833	102.
32	2.48	83.	1.40	74.3	2.691	107.
33	2.11	85.	1.49	81.	2.344	113.

Continued

j	1973		1975		1976	
	y_{1j}	w_{1j}	y_{2j}	w_{2j}	y_{3j}	w_{3j}
34	2.34	85.	1.56	81.3	1.286	134.
35	2.16	91.	1.34	81.3	1.141	132.
36	2.18	94.	1.59	90.6	3.541	143.
37	2.16	109.	1.93	91.9	2.122	175.
38	2.66	111.	1.82	92.	2.711	180.
39	2.76	113.	1.35	93.2	2.453	170.
40	3.0	123.	1.38	97.9	2.218	190.
41	3.0	128.	1.37	104.6	1.410	44.3
42	3.93	197.	1.91	106.6	1.131	54.6
43			1.64	107.	2.141	58.
44			2.38	109.7	1.693	65.
45			1.32	117.	1.323	43.8
46			1.89	117.7	1.736	54.3
47			1.81	135.9	1.705	74.2
48			1.51	138.8	1.568	77.1
49			1.72	147.7	2.503	86.2
50					.984	22.3
51					1.108	26.4
52					1.086	28.7
53					.976	34.5
54					1.812	145.6
55					1.402	39.4
56					1.832	168.
57					1.723	120.4
58					1.876	127.2
59					2.144	210.
60					2.068	190.
Means	2.0969	64.5238	1.3808	65.8245	1.7909	87.93
Overall means	$\bar{y}$ = 1.7429 ; $\bar{w}$ = 74.2464					

The analytical model can be written as follows:

$$y_{ij} = \mu + \theta_i + \beta_i w_{ij} + \epsilon_{ij} \qquad \begin{cases} i = 1, \ldots, 3 \\ j = 1, \ldots, n_i \end{cases}$$

and the ANOVA is reported in Table XII.

Table XII

ANOVA of the model $y_{ij} = \mu + \theta_i + \beta_i w_{ij} + \varepsilon_{ij}$ [DF]

$\beta_i \mid \mu, \theta_i$	MODELS	$\theta_i \mid \mu, \beta_i$
SS (μ) = 458.69276 [1]		
SS ($\theta_i \mid \mu$) = 11.82622 [2]		SS ($\beta_i \mid \mu$) = 8.921592 [3]
SS (μ, θ_i) = 470.51898 [3]		SS (μ, β_i) = 467.614352 [4]
SS ($\beta_i \mid \mu, \theta_i$) = 13.55092 [3]		SS ($\theta_i \mid \mu, \beta_i$) = 16.455548 [2]
SS ($\beta \mid \mu, \theta_i$) = 11.86722 [1]		
SS ($\beta_i - \beta \mid \mu, \theta_i$) = 1.68370 [2]		
Residual SS = 22.57095 [145] (MS = 0.15566)		
TOTAL: SS = 506.64085 [151]		

From Table XII it is now possible to test if the 3 regression lines have the same slopes, by:

$$F_{2,145} = \frac{1.68370}{2 \times 0.15566} = 5.41$$

which is much larger than the critical value at the confidence level of .99 so that we conclude that the slopes differ significantly. Analogously, it is seen that also the intercepts differ significantly.

The estimates of the regression lines reported in Figure 1 are for:

1973: $y_1 = 1.2846 + 0.01259\ w_1$

1975: $y_2 = 0.9928 + 0.00589\ w_2$

1976: $y_3 = 1.2744 + 0.00587\ w_3$

Comparing the regression lines of the years 1975 and 1976 (without considering 1973) it results (Table XIII) that the slopes do not differ significantly so that the two lines can be considered parallel.

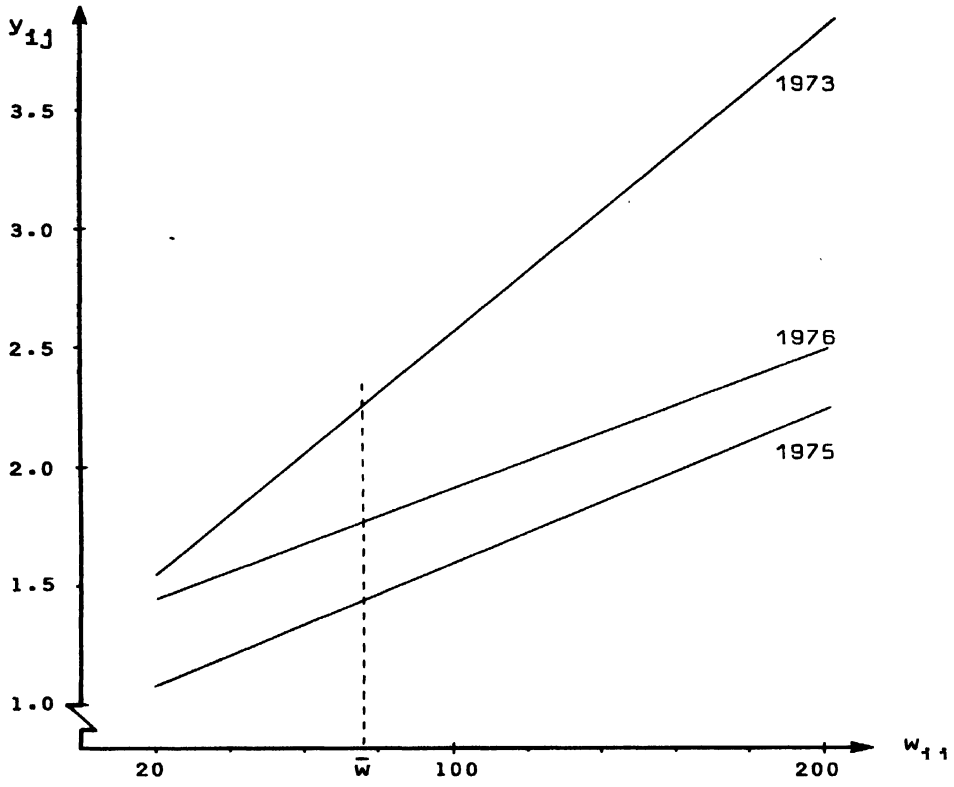


Figure 1 Regression lines of mercury content (white muscle) y_{ij} and body weight w_{ij}

Table XIII

ANOVA for the model $y_{ij} = \mu + \theta_i + \beta_i w_{ij} + \varepsilon_{ij}$ [DF]

$\beta_i \mid \mu, \theta_i$	MODELS	$\theta_i \mid \mu, \beta_i$

SS (μ) = 281.326906 [1]

SS $(\theta_i \mid \mu)$ = 4.536285 [1]	SS $(\beta_i \mid \mu)$ = 8.591268 [2]
SS (μ, θ_i) = 285.863192 [2]	SS (μ, β_i) = 289.918174 [3]
SS $(\beta_i \mid \mu, \theta_i)$ = 5.98745 [2]	SS $(\theta_i \mid \mu, \beta_i)$ = 1.932468 [1]
SS $(\beta \mid \mu, \theta_i)$ = 5.98743 [1]	
SS $(\beta_i - \beta \mid \mu, \theta_i)$ = 0.00002 [1]	

Residual	SS = 20.369698 [105]	(MS = 0.193997)
TOTAL:	SS = 312.22034 [109]	

Clearly, in such a situation, where the regression on the concomitant variable is so significant, it is almost useless to describe the results (of every year) by a mean and a standard deviation. At least the comparisons should be referred to a single point of the dependent variables, e.g. the mean $\bar{w}$, for which the regression lines assume the value of: $\bar{y}_i + b_i\ (\bar{w} - \bar{w}_i)$, with standard deviation (or error of prevision):

$$\frac{n+1}{n} \cdot \frac{(\bar{w} - \bar{w}_1)^2}{\sum_i (w_{ij} - \bar{w}_1)^2}$$

Since in this case $\bar{w} = 74.246$ we obtain, with s_ε, the square root of the residual MS (of Table XIII) the following values:

Years	1973	1975	1976
Expected value	2.219	1.430	1.711

with an error of prevision of $\approx .157$.

1.7.7 Example of a set-up of a (quality) control plan

We consider now the problem of the quality control of the determinations in an experiment like that of Table V. Further we will guess the most convenient number of times determinations must be performed according to some prescriptions.

We start analysing the ranges of the determinations of all samples and report them in Table XIV. We prefer an analysis based on ranges instead of standard deviations because the number of repetitions is as small as 2 (= m). Sample 1 because of repeated pretreatments is counted twice.

For each sample, at the bottom of Table XIV we have reported the sum of the ranges and their number q_i. The mean range is hence:

$$\bar{R} = \frac{414}{90} = 4.60$$

and the variance is estimated (see Chapter 2.4) by $\bar{R}/d_m = \bar{R}/d_2$ (d_2 from Table A-2). Hence an upper control limit (2σ) for the ranges is given by:

$$\bar{R}_{2\sigma} = \bar{R} + 2\,\bar{R}/d_2$$

$$= 4.60 + 2 \times 4.60 / 1.128$$

$$= 12.76 \ .$$

Table XIV

Ranges (× 1000) within determinations of samples i=1; 9 of Table V

j \ i	1		2	3	4	5	6	7	8	9		Σ
1	2	2	6	4	4	0	0	0	10	2	4	
2	2	3	6	15	0	6	1	7	2	0	1	
3	4	0	3	2	2	1	23	3	5	0	0	
4	5	3	24	0	3	1	16	3	7	3	4	
5	1	2	10	4	1	4	2	12	1	2	2	
6	0	0	11		1				0	0	1	
7	25	0	1		0				13	6	9	
8	18	0	6		2				3	7	5	
9	16	8	0		7				4	0	4	
10	0	6			4				15	7	1	
21											4	
q_i	10	10	9	5	10	5	5	5	10	21		90
$\sum R_{ij}$	73	24	67	25	24	12	42	25	60	62		414

It is seen from Table XIV that 9 determinations which are underlined lie above the 2σ control limit. The reasons for these exceptional ranges should be verified in the laboratory. The mean ratio of ranges above the upper control limit is: $\bar{p} = 9/90 = .1$, whilst the ratios of the single samples range from $p = 0$ to $p = .4$.

Given this situation of observed exceptional ranges (which could be different, if for instance 3σ upper limits had been chosen), let us try to find the most convenient number of determinations. Following the rules of 1.6 we have to state the precision $\varepsilon = \pi - p$ and the confidence level $1 - \alpha$.

In the precision statement π stands for a limiting acceptance level, which must be higher than the observed rate p, since this a typical unilateral decision statement. Hence $\varepsilon > 0$. In our case we could choose $\pi_1 = .15$ or $\pi_2 = .20$, etc. With the confidence level of $1 - \alpha = .95$, then analogously to (1.6-2) we can write:

$$z_{(1-\alpha)} = \frac{\pi - p}{\sqrt{\frac{\pi(1-\pi)}{n}\left(\frac{N-n}{N-1}\right)}} \quad ; \quad \pi > p$$

where z is the normal variate, $\pi(1-\pi)$ the expected variance of the number of exceptional ranges, N the total number of individuals, and $n \leq N$ the estimated number of individuals to be analysed in accordance with the preliminary statements. It follows:

$$1 + \frac{A}{N} \qquad \textit{with} \qquad A = \frac{z^2_{(1-\alpha)}\,\pi(1-\pi)}{(\pi - p)^2} \qquad \textit{and} \qquad \pi > p$$

For $\pi \leq p$ we will always leave $n = N$, what means that the sample equals the population.

Taking the q_1 values of Table V as population sizes N (besides ∞), we have evaluated, for $p = .1$ and some theoretical ratios π with $\alpha = .05$, the expected number n, obtaining:

	N = ∞	21	10	9	5
π = .15	n = 138	19	10	9	5
= .20	44	15	9	8	5
= .25	23	11	7	7	5
= .30	15	9	6	6	4

That means for instance: If we control at a confidence level of .95, the determinations of N individuals, such that the percentage of ranges falling above the 2σ limit is not higher than 25% (i.e. $\pi = .25$), then for $N = 21$ it is sufficient to repeat only 11 determinations, for $N = 10$ only 7, whereas for $N = 5$ all 5 determinations must be repeated, i.e. must be taken twice.

It will be noted that for a universe ($N = \infty$) the values of n reported in the above table are extraordinarily low as soon as $\pi \geq .20$.

1.7.8 Example of the determination of the sample sizes in an experiment with two samples

The determination of the sample size in an experiment with 2 samples.

Using data from Larsson (1975), $n_1 = 10$ and $n_2 = 9$ flounders (control and experimental respectively) were exposed during 4 and 9 weeks to 50 µg Cd/l (CS = 7%). At the end of each experiment the hematocrit % have been measured. The mean ± the standard error and the t-test (1.6-7) have been evaluated, obtaining following results:

	Control	50 µg/l	t_{17}
4 weeks	21.8 ± .9	18.8 ± .9	2.35
9 weeks	21.1 ± .8	16.5 ± .8	4.05
	$n_1 = 10$	$n_2 = 9$	

In the first experiment the significance level of the differences between control and 50 µg/l means lies between 2 and 5 percent, in the second below 1 percent. If the significance of the difference between 18.8 and 21.8 is retained to be too weak (hence the significance level too high) to decide that a dose of 50 µg Cd/l produces a real effect on the hematocrit % in four weeks, a lower significance level and hence a higher number of observations must be chosen. How must we choose n_2 with $n_1 = 10$, or, alternatively, $n_1 = n_2 = m$ if we require a significance level of 1%?

We start putting in (1.6-8) $t = t_{17(.995)} = 2.891$ and obtain $n_2 = 24$. But, in this case, we had $t = t_{32(.995)} = 2.75$. Reapplying (1.6-8) we obtain $n_2 = 18$ for which $t_{26(.995)} = 2.779$ and $n_2 = 19$. Since the next iteration would lead to a value between 18 and 19, $n_2 = 19$ could be assumed as the desired value. For $n_1 = n_2 = m$, using (1.6-9) we obtain after 3 iterations $m = 14$.

This means that, if a 1% significance level is required, with a given sample size of $n_1 = 10$ the sample size n_2 should at least be 19, or both samples should at least have 14 individuals.

2. TESTS FOR BIOLOGICAL POLLUTION RESEARCH

2.1 Introduction

2.1.1 Hypotheses, tests and errors

"The most frequent application of statistics in biological research is to test some scientific hypothesis. Statistical methods are important in biology because results of experiments are usually not clear cut and therefore need statistical tests to support decisions between alternative hypotheses. A statistical test examines a set of sample data and, on the basis of an expected distribution of the data, leads to a decision on whether to accept the hypothesis underlying the expected distribution (null hypothesis) or whether to reject that hypothesis and accept an alternative one" (Sokal and Rohlf, 1969). Hypotheses which are most frequently tested are those about:

(a) Frequencies, e.g. percentages of larvae hatching from eggs or fish being killed in a test;

(b) Location measures, e.g. mean (mode, median) body length;

(c) Measures of variability, e.g. variance or range of survival times, accummulation rates;

(d) Dependence (regression) and interdependence (correlation), e.g. doses (concentration) response relationship and correlated changes of oxygen uptake and activity following exposure to harmful substances, respectively.

(e) Detection of outliers, randomness.

Statistical testing can be done by parametric and non-parametric methods. With the former, hypotheses are stated about values of the parameters in the population density function, with the latter not. Many of the non-parametric methods are further distribution free, what means that not even hypotheses about the form of the density function must be considered. Most distribution free tests can be derived using simple combinatorial formulas, so that each test has its own rationale.

For the parametric testing the rationale is almost the same for all tests and can be summarized as follows: First of all the population parameter θ, like the mean μ or the variance σ^2, has to be defined, generally with the density function of the population. We are then concerned with deciding whether some preassigned value θ_0 is acceptable in the light of the observations (= sample).

We thus have the null hypothesis $H_0: \theta = \theta_0$.

The alternative hypotheses can be very different. We can have a simple alternative hypothesis, whether the population parameter is θ_1 instead of θ_0, what we write $H_1: \theta = \theta_1$. Or we can have composite alternative hypotheses, what means that the population parameter is larger or smaller than θ_0. In this case we write $H_1: \theta > \theta_0$ or $H_1: \theta < \theta_0$. These hypotheses are unilateral, whereas the composite alternative hypothesis that θ is different from θ_0, i.e. $H_1: \theta \neq \theta_0$, is bilateral. In this paper we refer only to composite alternative hypotheses.

The probability to refuse the null hypothesis, i.e. to refuse H_0, is measured by the confidence level $1-\alpha$ (α, small, positive value, usually $\alpha = .05$, .01 or .001). 100 α is said to be the significance level of the test, and α the size of the test or the type I error.

In the case of a simple alternative hypothesis the probability to accept H_0 when it is false, i.e. when the alternative hypothesis is true, is measured by β, the type II error.

The probability $1-\beta$ represents the power of the test to detect a false null hypothesis

The decision process in statistical inference can be summarized as follows (Conover, 1971):

THE DECISION:

The true situation		ACCEPT H_0	REJECT H_0
	H_0 is true	Correct decision Probability = $1-\alpha$	Type I error Probability = α (level of significance)
	H_0 is false	Type II error Probability = β	Correct decision Probability = $1-\beta$ (power of the test)

To test any hypothesis on the basis of a sample of observations, we must divide the sample space, i.e. all possible sets of observations, into two regions: the acceptance region where H_0 will be accepted at the given confidence level and the critical region, complementary to the first, where it will be refused. In the case of unilateral hypotheses one critical value divides the acceptance from the critical region, whereas for bilateral or two-tailed hypotheses there are two critical values, an upper and a lower limit.

In Figure 2 we have reported the critical value $z_{(.95)}$ at the confidence level of $1-\alpha = .95$ to test the unilateral hypothesis whether an observed value z can be considered belongir to a normal population with mean μ_0, or whether we must suppose that it derives from a normal population with a larger mean value.

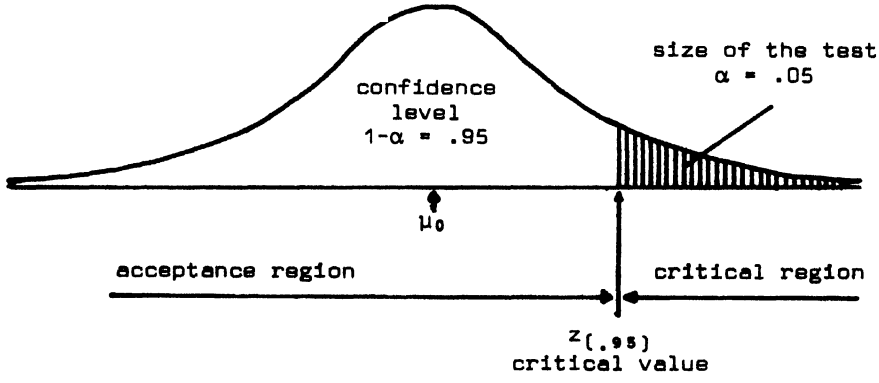


Figure 2 Example of a unilateral hypothesis

The hypothesis to be tested is hence $H_0: \mu = \mu_0$ against the alternative $H_1: \mu > \mu_0$ If $z < z_{(.95)}$ we accept the null hypothesis H_0, otherwise we refuse it.

In Figure 3 we consider the test of the bilateral hypothesis whether an observed value can be considered belonging to a normal population with mean μ_0, or whether we must suppose that it derives from a normal population whose mean is different (hence larger or smaller) from μ_0. This hypothesis will then be treated as two unilateral hypotheses at sizes of the test $\frac{\alpha}{2}$:

$$\text{(i)} \quad H_0: \mu = \mu_0 \quad ; \quad H_1: \mu > \mu_0$$

$$\text{(ii)} \quad H_0: \mu = \mu_0 \quad ; \quad H_1: \mu < \mu_0 .$$

For the first, the critical value $z_{(1-\frac{\alpha}{2})}$ is obtained and for the second $z_{(\frac{\alpha}{2})}$. Hence the bilateral hypothesis:

$$H_0: \mu = \mu_0 \quad ; \quad H_1: \mu \neq 0$$

which resumes both unilateral hypotheses will be accepted if the observed z lies within $z_{(\frac{\alpha}{2})}$ and $z_{(1-\frac{\alpha}{2})}$, hence if $z_{(\frac{\alpha}{2})} < z < z_{(1-\frac{\alpha}{2})}$.

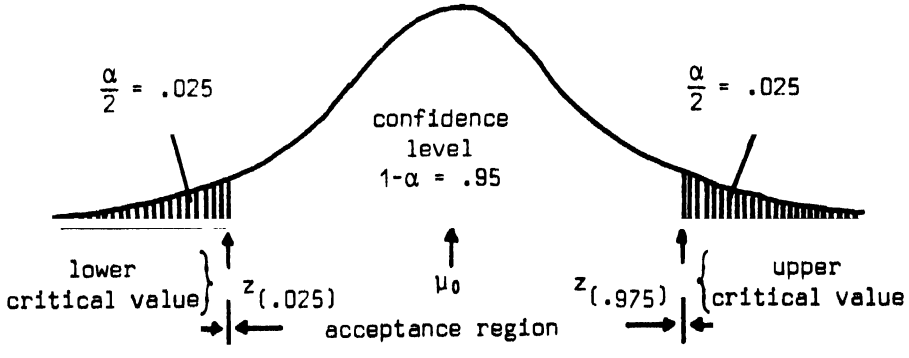


Figure 3 Example of a bilateral hypothesis

Since the critical values of bilateral hypotheses coincide with the correspondent unilateral critical values with the half significance level, in the following we will not treat unilateral and bilateral hypotheses separately, since only the significance level has to be adjusted.

We will now describe a number of tests following the order of the 5 points, (a)...(e). Examples and applications can be found in Section 2.7. Some other examples, especially about analysis of variance and regression, are in Section 1.7.

The number of tests described hereafter is obviously limited and cannot pretend to cover all possible hypotheses, but should enable the research worker to find easily the mos important tests concerning his experiments.

There are many books about biometrics on a more or less mathematically basis like Rao (1952), Snedecor (1956), Sokal and Rohlf (1969), in which almost all tests are described and the necessary tables with the critical values are reported. Further classical table works are of Fisher and Yates (1963), Pearson and Hartley (1958), Diem (1962) and Stange and Henning (1966).

Because of their repeated use we will always write the standardized normal variate (in unities of the standard deviation) as z:

$$z = \frac{x - \mu}{\sigma} ,$$

Student's t-variate with ν degrees of freedom (DF):

$$t_\nu = \frac{x - \mu}{s} \sqrt{n} ,$$

the chi-square variate with ν DF by:

$$\chi^2_\nu = \frac{(n-1)s^2}{\sigma^2} ,$$

and the variance ratio variate with ν_1 DF for the variance s_1^2 and ν_2 DF for s_2^2 by:

$$F_{\nu_1,\nu_2} = \frac{s_1^2}{s_2^2} .$$

The letter n indicates always the sample size, i.e. the number of all observations.

Critical values of z, t, χ^2 and F and all other tests proposed, at the confidence level of $1-\alpha$, will be indicated by the letter of the test, followed by the subscript within parentheses indicating the confidence level:

$$z_{(1-\alpha)}, \quad t_{\nu(1-\alpha)}, \quad \chi^2_{\nu(1-\alpha)}, \quad F_{\nu_1,\nu_2(1-\alpha)} .$$

Tables of the critical values of z are reported in Table A-3[1], of t in Table A-4, of χ^2 in Table A-5 and of F in Table A-6. If the value of a function is searched for an argument (not reported in the table) which is intermediate between two arguments reported in the table, it can be guessed by the method of interpolation. The correct way to find interpolated values of the function is described at the end of the table annex.

Tables A-1 to A-25 are placed in a special table annex at the end of the paper

2.1.2 Confidence intervals and tests

It is known (Chapter 1.3) that any interval which specifies two statistics (estimates) L_1, L_2 such that the probability that a population parameter θ lies within the interval (L_1, L_2) is $1-\alpha$, where α is a preassigned positive constant (< 1), then the interval (L_1, L_2) is said to provide an interval estimate, or confidence interval for the parameter θ at the confidence level $100(1-\alpha)\%$. This is generally indicated by:

$$P\,[L_1 < \theta < L_2] \;=\; 1-\alpha\,. \qquad (2.1\text{-}1)$$

Often the sampling distribution of a function of the population parameter, like a statistical test, $T(\theta)$, is known. Determining two statistics T_1 and T_2 (2.1-1) can be transformed to:

$$P\,[T_1 < T(\theta) < T_2] \;=\; \int_{T_1}^{T_2} f(T)\,d\,T \;=\; 1-\alpha\,. \qquad (2.1\text{-}2)$$

Solving $T(\theta)$ for θ and adapting T_1 and T_2 the confidence limits of θ as functions of T_1 and T_2 can be found. Suppose that T has been chosen to test the hypothesis:

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_1 : \theta \neq \theta_0\,.$$

Then chosing T_1 and T_2, under the null hypothesis, at the $100(1-\alpha)$ confidence level, the $100(1-\alpha)$ confidence interval for θ_0: $[L_1(T_1), L_2(T_2)]$ can be obtained. It follows that the values of θ within the confidence interval, $L_1(T_1) < \theta < L_2(T_2)$ would yield a value of the T-test within the acceptance region(T_1, T_2). In other words, if $\hat{\theta}$ lies in its $100(1-\alpha)$ confidence interval, test T will not refuse H_0 at the significance level of α, and vice-versa. Also, the comparison of 2 parameters like $H_0: \theta_1 = \theta_2$ can be obtained by the use of confidence intervals of both parameters. In fact,

(i) If no point is common to both intervals, H_0 can be refused at a smaller significance level than α;

(ii) If both parameters lie within the limits of the confidence interval of the other, then H_0 can be accepted ;

(iii) If the intervals overlap in some way, other than under (ii), then multiplying the interval of the sample of size n_1 by:

$$\frac{1}{2}\sqrt{\frac{n_1+n_2}{n_2}}$$

and the interval of the sample of size n_2 by: $\frac{1}{2}\sqrt{\frac{n_1+n_2}{n_1}}$ we can conclude with good approximation that no over-lapping leads to the refusal of H_0, whilst over-lapping to the acceptance of H_0. It may be observed that for $n_1 \approx n_2$ both multipliers reduce to $\sim \frac{1}{\sqrt{2}}$.

It is thus seen that for both tests it is easy to obtain from statistical tests confidence intervals and vice-versa. Here we shall limit ourselves to hypotheses testing.

2.2 Hypotheses about frequencies

2.2.1 Test whether the frequency π of a character in a population equals π_0 or it is larger

The hypothesis to be tested is:

$$H_0 : \pi = \pi_0 \quad \text{with the alternative} \quad H_1 : \pi > \pi_0 . \tag{2.2-1}$$

Sometimes H_0 is put in the more general form $H_0 : \pi \leq \pi_0$.

If k observations in a sample of size n have the examined character, whose frequency in the population is hypothesized to be π_0, the test statistic is given by:*

$$F_{\nu_1,\nu_2} = \frac{1-\pi_0}{\pi_0} \cdot \frac{k}{n-k+1} \quad ; \quad \nu_1 = 2(n-k+1), \quad \nu_2 = 2k . \tag{2.2-2}$$

If the observed F is larger than the tabulated (critical) F, hypothesis (2.2-1) must be rejected.

For $\pi_0 \geq \frac{12}{12+n}$ test (2.2-2) can be approximated, putting $p = \frac{k}{n}$, by the normal variate:

$$z = \frac{p - 1/(2n) - \pi_0}{\sqrt{\pi_0(1-\pi_0)}} \sqrt{n} . \tag{2.2-3}$$

H_0 is rejected if $z \geq z_{(1-\alpha)}$. For large n the correction factor $\frac{1}{2n}$ could be eliminated.

2.2.2 Test whether the frequency π of a character in a population equals π_0 or it is smaller

The hypothesis to be tested is:

$$H_0 : \pi = \pi_0 \quad \text{with the alternative} \quad H_1 : \pi < \pi_0 . \tag{2.2-4}$$

The test statistic is now:*

$$F_{\nu_1,\nu_2} = \frac{\pi_0}{1-\pi_0} \frac{n-k}{k+1} \quad ; \quad \nu_1 = 2(k+1), \quad \nu_2 = 2(n-k) . \tag{2.2-5}$$

H_0 is rejected if $F_{\nu_1,\nu_2} \geq F_{\nu_1,\nu_2(1-\alpha)}$.

For $\pi_0 \leq \frac{n}{12+n}$ test (2.2-5) can be approximated analogously to (2.2-3) by:

$$z = \frac{\pi_0 - p - 1/(2n)}{\sqrt{\pi_0(1-\pi_0)}} \sqrt{n} . \tag{2.2-6}$$

H_0 is rejected if $z \geq z_{(1-\alpha)}$.

---

*Approximation of the binomial test

2.2.3 Comparison of 2 frequencies: test of independence (both margins fixed)

Suppose we have two samples of sizes n_1 and n_2, $n_1 + n_2 = n$, with k_1 and k_2 observations which detain the same or different characters. If they refer to the same character, the question is whether the samples belong to the same population. If they refer to different characters, the question is whether the frequencies of the characters in the respective populations are the same. In both cases the hypothesis to be tested can be written as follows:

$$H_0 : \pi_1 = \pi_2 = \pi \quad \text{against} \quad H_1 : \pi_i \neq \pi \; ; \; i = 1,2 \tag{2.2-7}$$

To find the test statistic we build the four-fold table with the margins:

	Sample 1	Sample 2	Sums
Number of observations with the character	k	k_2	$k_1 + k_2$
Number of observations without the character	$n_1 - k_1$	$n_2 - k_2$	$n - k$
Sums	n		

Now we individuate the smallest entry in the four-fold table, say k_1, and compute the probabilities*:

$$P_i = \frac{k!\,(n-k)!\,n_1!\,n_2!}{i!\,(n_1-i)!\,(k-i)!\,(n-n_1-k+i)!\,n!} \; ; \; i = 0 \ldots k_1 \tag{2.2-8}$$

and add all probabilities to give:

$$P = \sum_{i=0}^{k_1} P_i \; . \tag{2.2-9}$$

This is the exact test of Fisher (1954) and gives directly the type I error. If $P \leq \frac{\alpha}{2}$, H_0 is rejected. It should be observed that this procedure is only correct if (Müller, 1961):

$$\frac{k_1\,(n_2-k_2)}{(k_2+1)\,(n_1-k_1+1)} \leq 1 \; .$$

The computations of this test are somewhat long, depending on the size of the sample. When the lowest cell frequency in the table is greater than 10, the method is quite impractical. When no marginal total is greater than 15, Biometrical tables (Pearson and Hartley, 1958) give directly the probabilities P, and for marginals not greater than 30 we can find them in Diem (1962).

* $k! = 1 \times 2 \times 3 \times \ldots \times k$. See Table A-1

For values of $n > 20$ and no entry of the four-fold table smaller than 5, or for values of $n > 40$ and no empty entry, then H_0 (2.2-7) can be tested by:

$$\chi_1^2 = \frac{[k_1(n_2-k_2) - k_2(n_1-k_1)]^2 n}{n_1 \cdot n_2 \cdot k \cdot (n-k)} . \qquad (2.2\text{-}10)$$

If $\chi_1^2 \geq \chi^2_{1(1-\alpha)}$, H_0 is rejected.

If c 4-fold tables must be tested for independence, then the joint test:

$$\chi_c^2 = \sum_{j=1}^{c} \chi_{1,j}^2$$

can be computed, comparing it, like (2.2-10) with $\chi^2_{c(1-\alpha)}$.

Another test statistic to prove H_0 is based on the arc sin transformation (in radiants):

$$\frac{| \text{arc sin } \sqrt{p_1} - \text{arc sin } \sqrt{p_2} |}{\frac{1}{2}\sqrt{\frac{n_1 + n}{n_1 \, n_2}}} ; \quad p_1 = \frac{k_1}{n_1}, \quad p_2 = \frac{k_2}{n_2}. \qquad (2.2\text{-}11)$$

If $z \geq z_{(1-\frac{\alpha}{2})}$, H_0 is rejected.

2.2.4 Comparison of c frequencies: test of homogeneity (one margin fixed)

We consider now a population whose elements differ for r (≡ rows) characters, with frequencies π_j, $j = 1 \ldots r$. Consider c (≡ columns) samples of size $n_{i\cdot}$, $i = 1 \ldots c$. For the i-th sample there are n_{ij}, $\sum_j n_{ij} = n_{i\cdot}$, individuals which posses character j. Thus we can build the two-way cross table:

	Samples					
characters	1	...	i	...	c	sums
1	n_{11}	...	n_{i1}	...	n_{c1}	$n_{\cdot 1}$
⋮	⋮		⋮		⋮	⋮
j	n_{1j}	...	n_{ij}	...	n_{cj}	$n_{\cdot j}$
⋮	⋮		⋮		⋮	⋮
r	n_{1r}	...	n_{ir}	...	n_{cr}	$n_{\cdot r}$
sums	$n_{1\cdot}$	...	$n_{i\cdot}$	...	$n_{c\cdot}$	n

(2.2-12)

The hypothesis to test is:

$$H_0: \pi_{ij} = \pi_j \text{ , all } i \text{ , } j \text{ against } H_1: \pi_{ij} \neq \pi_j \text{ some } i,j \text{ .} \qquad (2.2\text{-}13)$$

If H_0 is true, the frequencies π_j are estimated by $p_j = n_{\cdot j}/n$.
The following conditions must be satisfied:

1) No expected value $n_i p_j = \frac{n_{i\cdot}\, n_{\cdot j}}{n}$ must be smaller than 1 .

2) Only 20% of the above expected values must be smaller than 5 .

If these conditions are not fulfilled some columns and/or rows of (2.2-12) can be joined so as to meet the conditions.

Then H_0 can be tested by:

$$\chi^2_{(c-1)(r-1)} = n\left(\sum\sum \frac{n_{ij}^2}{n_{i\cdot}\, n_{\cdot j}} - 1\right) . \qquad (2.2\text{-}14)$$

If $\chi^2 \geq \chi^2_{(1-\alpha)}$, H_0 is rejected.

This test can also be used to test independence. For a 4-fold table it reduces to (2.2-10). If the population frequencies would be known or predetermined, H_0 can be tested by:

$$\chi^2_{c(r-1)} = \sum\sum \frac{(n_{ij} - n_i\, \pi_j)^2}{n_i\, \pi_j} . \qquad (2.2\text{-}15)$$

This corresponds to the test of goodness of fit (2.2-18) for the comparison of c observed frequency distributions with one theoretical. Thus if H_0 can be accepted, the c distributions are independent.

If χ^2 of (2.2-15) is greater than the correspondent $\chi^2_{(1-\alpha)}$, H_0 is refused. Hypothesis (2.2-13) can also be tested by Kullback's test (Kullback, 1959), which is also known as G-test (Sokal and Rohlf, 1969):

$$G_{(c-1)(r-1)} \simeq \chi^2_{(c-1)(r-1)} \qquad (2.2\text{-}16)$$

$$= 2\Big(\sum\sum n_{ij} \ln n_{ij} - \sum n_{i\cdot} \ln n_{i\cdot} - \sum n_{\cdot j} \ln n_{\cdot j} + n \ln n\Big)$$

where ln indicates the natural logarithm (= $\log_e$).

If $G \geq \chi^2_{(1-\alpha)}$, H_0 is rejected. The utility of the G-test lies especially in its application to more complex designs.

The case of no margins fixed is treated in 2.2.7.

2.2.5 Comparison of an observed with a theoretical distribution: tests of goodness of fit

A discrete or continuous population density function $\psi(x)$ is predesignated $(= \psi_0(x))$. The total range of variability of x is sub-divided into r classes, so that for each class the frequencies π_j, $j = 1 \cdots r$, can be evaluated. Now a sample of size n is considered, with n_j observations for each class, $\Sigma n_j = n$. To test the hypothesis,

$$H_0 : \psi(x) = \psi_0(x) \quad \text{against} \quad H_1 : \psi(x) \neq \psi_0(x) \tag{2.2-17}$$

we must first evaluate the expected frequencies $n\pi_j$ and check the following conditions:

(1) If $r = 2$ both $n\pi_1$ and $n\pi_2$ must be greater than 5,

(2) If $r > 2$ see conditions 1) and 2) of Section 2.2.4

The null hypothesis will be rejected if the goodness of fit test

$$\chi^2_{r-1} = \sum \frac{(n_j - n\pi_j)^2}{n\pi_j} \geq \chi^2_{r-1(1-\alpha)} . \tag{2.2-18}$$

The extension of this test to c distributions is given by (2.2-15).

If the expected frequencies have been obtained through an estimate of m parameters, then the DF of χ^2 must be reduced by further m.

The same null hypothesis can be tested by the Kolmogoroff-Smirnov test (Kolmogoroff, 1941; Smirnov, 1948) for continuous variates and for samples with more than 5 observations. All we have to do is to write down in a column the cumulated population frequencies: $\Phi(x_1) = \pi_1$, $\Phi(x_2) = \pi_1 + \pi_2 \cdots$ and in general $\Phi(x_j) = \sum_{h=1}^{j} \pi_h$. Correspondingly we evaluate the cumulated relative frequencies of the ordered sample $(x_1 < x_2 \cdots < x_r)$

$$F(x_j) = \frac{1}{n} \sum_{h=1}^{j} n_h$$

and write them in a second column. In the third column we write the differences: $D(x_j) = \Phi(x_j) - F(x_j)$ and we select the maximum value of these differences without considering the sign:

$$D_n \quad \max | D(x_j) | . \tag{2.2-19}$$

If $D_n \geq D_{n(1-\frac{\alpha}{2})}$, H_0 is rejected (Table A-7). Unilateral hypotheses can also be tested. For large n $D_{n(1-\alpha)}$ can be approximated after Smirnow (1948):

$1-\alpha$	.95	.975	.99	.995
$D_{n(1-\alpha)}$	$\frac{1.22}{\sqrt{n+1}}$	$\frac{1.36}{\sqrt{n+1}}$	$\frac{1.52}{\sqrt{n+1}}$	$\frac{1.63}{\sqrt{n+1}}$

(2.2-20)

For discrete variables the above test procedure can be applied, with the only difference that now the confidence level is not equal to $1-\alpha$ but it is larger.

The Kolmogoroff-Smirnov test can also be generalized to include the alternative hypotheses:

$$H_1: \quad \Phi(x) > \Phi(x_0) \qquad (2.2\text{-}21)$$

and

$$H_1: \quad \Phi(x) < \Phi(x_0) \qquad (2.2\text{-}22)$$

In the first case the test variate is the maximum positive difference: $D^+(x_j)$; in the second the maximum negative difference in absolute value: $-D^-(x_j)$, which can be compared with the critical values D_n at the confidence level of $1-\alpha$.

2.2.6 Comparison of an observed with the normal distribution

2.2.6.1 Probit analysis

A widely used method of comparing an observed distribution function*, $F(x_j)$, with a normal one, $\Phi(x_j)$, is that of the probit transformation of the observed cumulated frequencies. If the observed values of x_j are normally distributed, the resulting probit transformation will be a straight line.

The procedure generally followed (Finney, 1971), given x_j and the cumulated frequencies $F(x_j)$, is to transform the latter, recurring to Table A-19, into empirical probits $Y(x_j)$. A straight line is then plotted through the empirical probits, either by eye, or by the usual regression methods, weighting the probits with the weights $w[Y(x_j)] = w_j$, which can be found in Table A-20. From the regression line the expected probits $Y'(x_j)$ can be guessed. For most purposes, these will be sufficient. For more precise results the expected probits will be corrected to working probits $y(x_j)$ by the transformations:

$$y(x_j) = Y_1 - [1 - F(x_j)]A \qquad \text{for} \qquad Y'(x_j) \geq 5$$

and (2.2-23)

$$y(x_j) = Y_0 + F(x_j)A \qquad \text{for} \qquad Y'(x_j) \leq 5 \; .$$

Here Y_0 and Y_1 are the minimum working and maximum working probits respectively, and A is the range. Y_0, Y_1 and A can also be found in Table A-20. A new weighted regression line (with weights w_j) is plotted through the working probits obtaining the new expected probits $Y''(x_j)$. If the correspondence between $Y''(x_j)$ and the expected probits used in the first cycle, $Y'(x_j)$, is not considered to be close enough, computations must be repeated starting from $Y''(x_j)$ until the desired degree of correspondence is obtained.

2.2.6.2 Normality tests

Besides the tests previously described in Section 2.2.5 and the tests for outliers in Section 2.6.1, other tests are used to test the hypothesis whether a distribution is consistent with the hypothesis of normality:

$$H_0: \quad \Phi(x) = \text{normal distribution.} \qquad (2.2\text{-}24)$$

*A distribution function is defined by the cumulated frequencies of a frequency distribution

In the following we describe two methods for testing the above hypothesis:

(i) The tests of skewness and kurtosis. We first evaluate

$$b_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\sqrt{n}\,\sum(x_i-\bar{x})^3}{[\sum(x_i-\bar{x})^2]^{3/2}} \qquad (2.2\text{-}25)$$

which is known as skewness, then

$$b_2 = \frac{\mu_4}{\mu_2^2} = \frac{n\,\sum(x_i-\bar{x})^4}{[\sum(x_i-\bar{x})^2]^2} \qquad (2.2\text{-}26)$$

which is known as kurtosis.

The first test, b_1, is sensitive to departures from normality in symmetry, hence with the alternative hypothesis H_1: $\Phi(x)$ = not normal and skew distribution. The second, b_2, is sensitive to departures from normality in the degree of peakedness of the distributions, hence with the alternative hypotheses H_1: $\Phi(x)$ = plati- or leptokurtic distributions.

For a normal distribution (symmetric and mesokurtic) we have $b_1 = 0$ and $b_2 = 3$. Positive values of b_1 indicate a longer right tail, negative a longer left tail of the central maximum of the distribution. Values of b_2 larger than 3 indicate leptokurtic distributions, with higher peak and over-dispersion; smaller than 3, but always $> 1 + b_1$, indicate platikurtic and under-dispersion when compared with the normal.

Approximated critical values for $|b_1|$ and b_2 at the 1% and 5% significance level, have been evaluated by Pearson (1930) and D'Agostino and Pearson (1973) from which we have taken the lines of Figure 5, whereas those of Figure 4 have been evaluated on the base of the sampling variance of b_1. The acceptance areas resulting from the combined decision $|b_1| \leq b_{1(\alpha)}$ and $b_{2(\alpha)}$ lower limit $\leq b_2 \leq b_{2(\alpha)}$ upper limit, lie almost completely within the contours evaluated by Bowman and Shenton (1975), but comprises those of Shenton and Bowman (1977), at least for the example reported by them.

Values of $|b_1|$ and b_2 falling within the acceptance region, permit thus to accept the null hypothesis (2.2-24) at the given confidence levels.

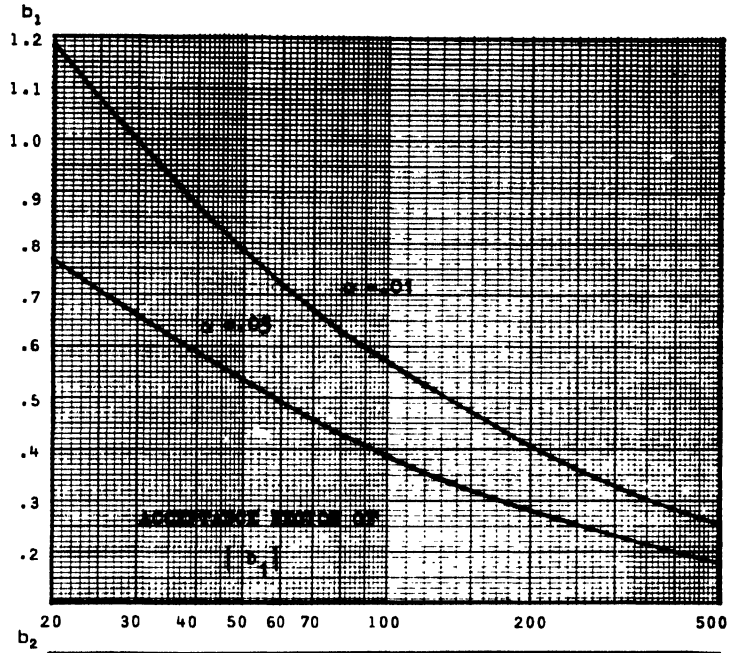


Figure 4 Acceptance region at the 1% and 5% significance level for the 3rd standardized moment b_1

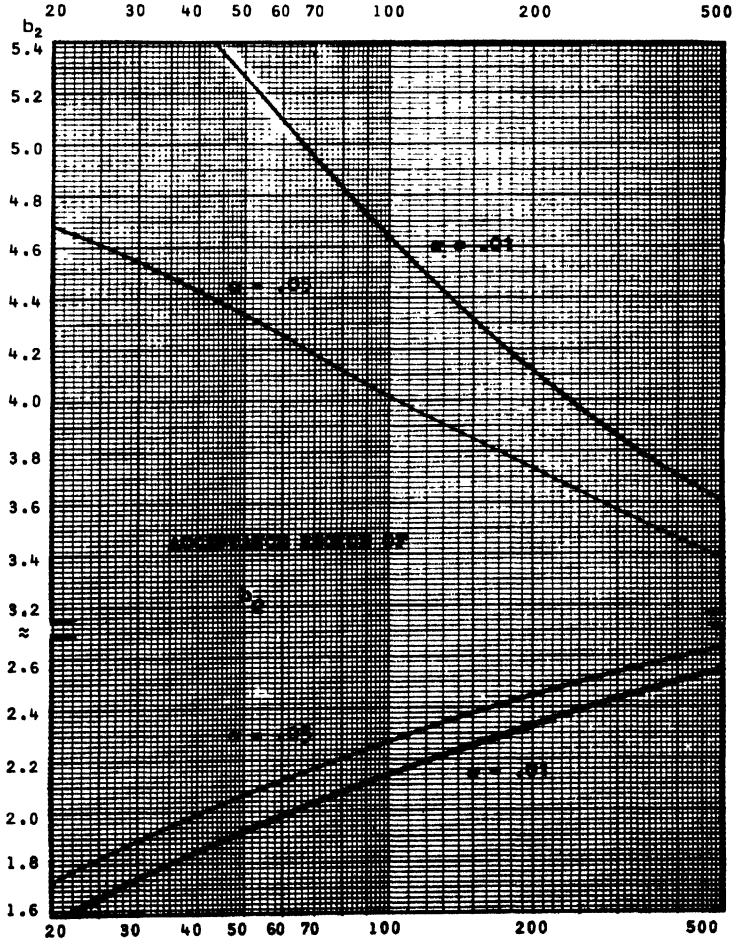


Figure 5 Bilateral acceptance region at the 1% and 5% significance level for the 4th standardized moment b_2

(ii) Generally more efficient tests, especially for low values of n (Shapiro et al., 1968) are those based on the ordered sequences $x_{(i)} > x_{(i-1)}$ as the W_n-test (Shapiro and Wilk, 1965) or as the normal plot correlation coefficient r_n (Filliben, 1965). The latter is, for small n, poorer than W for shorter tailed and skew alternatives, but marginally better than W for symmetric larger tailed alternatives. Further, the latter has the great advantage that it is completely computer implementable, what means that no coefficients need be stored.

To compute the value of W, given a complete random sample of size n, one first has to compute:

$$b = a_1 [x_{(n)} - x_{(1)}] + a_2 [x_{(n-1)} - x_{(2)}] + \cdots \qquad (2.2\text{-}27)$$

extending the sum for all positive differences within [], where the values of a_i are given in Table A-23. Then the W-test of hypothesis (2.2-24) is given by:

$$W_n = \frac{b^2}{(n-1)\, s^2} \qquad (2.2\text{-}28)$$

where s^2 is the sample variance. If $W_n \leq W_{n(\alpha)}$ (Table A-24), H_0 must be rejected.

To evaluate the normal plot correlation coefficient r_n, first the uniform order statistic medians m_i can be evaluated (with 3 decimal places) using

$$m_i = \begin{cases} 1 - .5^{1/n} & i = 1 \\ (i - .3175) / (n + .365) & i = 2, \ldots, n-1 \\ .5^{1/n} & i = n \end{cases} \qquad (2.2\text{-}29)$$

and finally the normal order statistic medians M_i, as computed from $\Phi^{-1}(m_i)$ or the inverse normal function (Table A-3) will be obtained. Hence evaluate the ordinary correlation coefficient:

$$r_n = \frac{\sum x_{(i)}\, M_i}{\sqrt{(n-1)s^2 \sum M_i^2}} \qquad (2.2\text{-}30)$$

with critical values reported in Table A-25. If $r_n \leq r_{n(\alpha)}$, H_0 (2.2-24) must be rejected.

2.2.7 Comparison of two distributions

Consider a sample of size n_1 from a population with distribution function $\Phi_1(x)$, (Section 2.2.5) and a sample of size n_2 from a population with distribution function $\Phi_2(x)$, where x is a continuous variate. We test the hypothesis whether both samples derive from the same population, or equivalently if both distribution functions are equal:

$$H_0: \Phi_1(x) = \Phi_2(x) = \Phi(x). \qquad (2.2\text{-}31)$$

First we transform the n_1 and n_2 observed values in one ranked sequence, i.e. we associate rank 1 to the smallest value of the n_1 and n_2 observed, rank 2 to the next and so on. The rank of the highest value is then n_1+n_2. If some values are equal, the mean rank is associated to them. Now we choose the smaller sample, say the first if $n_1 \leq n_2$, and evaluate the sum of the ranks R_1 corresponding to the values of that sample, then compute the Mann-Whitney-Wilcoxon tests:

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \quad ; \quad U_2 = n_1 n_2 - U_1 . \qquad (2.2\text{-}32)$$

If the alternative hypothesis is $H_1: \Phi_1(x) > \Phi_2(x)$, H_0 will be rejected if $U_2 \leq U_{n_1,n_2(1-\alpha)}$.
If the alternative hypothesis is $H_1: \Phi_1(x) < \Phi_2(x)$, H_0 will be rejected if $U_1 \leq U_{n_1,n_2(1-\alpha)}$.
If the alternative hypothesis is $H_1: \Phi_1(x) \neq \Phi_2(x)$, H_0 will be rejected if the smaller of U_1 and U_2 is smaller than $U_{n_1,n_2(1-\frac{\alpha}{2})}$. Critical values are reported in Table A-8. If n_1 or $n_2 > 20$, or $n_1+n_2 > 40$, we can recur to the normal approximation:

$$U_{n_1,n_2(1-\alpha)} = \frac{n_1 n_2}{2} - z_{(1-\alpha)} \sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}} . \qquad (2.2\text{-}33)$$

Another way of testing hypothesis (2.2-31), if $|n_1-n_2| \leq 5$ and $n = n_1+n_2 \leq 50$ is the X-test of Van der Waerden (1971). This consists in transforming the ranks of one of the two samples, say $r_1 \cdots r_{n_1}$ into fractions

$$\frac{r_1}{n+1}, \quad \frac{r_2}{n+1}, \ldots, \frac{r_{n_1}}{n+1} .$$

If these are considered as the cumulative frequencies of a distribution, $F(\frac{r_j}{n+1})$, we can transform them into normal variates, $Y^{-1}(x_j)$, by inspection of Table A-3. Then we form Van der Waerden's test:

$$X_n = \left| \sum_j Y^{-1}(x_j) \right| . \qquad (2.2\text{-}34)$$

If $X_n \geq X_{n(1-\alpha)}$ (Table A-21), the null hypothesis must be rejected. Since the test is symmetric, unilateral hypotheses can also be tested easily.

Still another way is given to test the same hypotheses. This is done dichotomising the values of both samples in the ranked sequence, writing the letter A for all values of the first and B for all values of the second sample, the ranked sequence transforms to a sequence of A and B. A sequence of one or more like elements preceeded and/or followed by unlike elements is called a run. The original sequence of x_i is thus transformed into a sequence of r runs. The null hypothesis (2.2-31) that both samples derive from the same population can then be tested by inspection of Table A-9, where the lower and upper limits r_1, r_2, of the acceptance interval are given at the confidence level of 1-α. If $r_1 < r < r_2$, H_0 can be accepted, otherwise rejected.

If one of the two sample sizes is greater than 20, a normal approximation can be used:

$$- \frac{r - [2\, n_1 n_2 \,/\, (n_1+n_2) - 1]}{\sqrt{[2\, n_1 n_2\, (2n_1 n_2 - n_1 - n_2)] \,/\, [(n_1+n_2)^2 (n_1+n_2-1)]}} \qquad (2.2\text{-}35)$$

which reduces for $n_1 = n_2 = c$ to:

$$z = \frac{r - c - 1}{\sqrt{n(n-1) \,/\, (2n-1)}} \,. \qquad (2.2\text{-}36)$$

If $z_{(\frac{\alpha}{2})} < z < z_{(1-\frac{\alpha}{2})}$ The null hypothesis can be accepted.

The above hypothesis (2.2-31) can also be tested by the Kolmogoroff-Smirnov test (2.2-19) if the x_i classes are the same, with critical values:

$$\lambda_{(1-\alpha)} \sqrt{\frac{n_1+n_2}{n_1\, n_2}} \qquad (2.2\text{-}37)$$

with $\lambda_{(1-\alpha)}$ given by Stange and Henning (1966), after Smirnow (1948):

$1-\alpha$	.95	.975	.99	.995
$\lambda_{(1-\alpha)}$	1.36	1.48	1.63	1.73

If the two samples are linked, when for one reason or another the observations have been obtained pair-wise, then hypothesis (2.2-31) can be tested by Wilcoxon's pair difference test. In this case obviously $n_1 = n_2$. We first calculate all differences $d_j = x_{1j} - x_{2j}$. Suppose the non-null differences are n_d, n_d should not be much smaller than n. Then we transform the absolute values of the n_d differences, $|d_j|$, in a ranked sequence.

Now the ranks of the positive differences are added to give T^+ and the sum of negatives are added to give T^-, $(T^+ + T^- = \frac{1}{2}\, n_d(n_d + 1))$. The null hypothesis will be rejected for:

$$\begin{aligned} &H_1 : \Phi_1(x) > \Phi_2(x) \quad \text{if} \quad T^+ \leq T_{n_d(1-\alpha)} \\ &H_1 : \Phi_1(x) < \Phi_2(x) \quad \text{if} \quad T^- \leq T_{n_d(1-\alpha)} \\ &H_1 : \Phi_1(x) \neq \Phi_2(x) \quad \text{if} \quad \min(T^+, T^-) \leq T_{n_d(1-\frac{\alpha}{2})} \,. \end{aligned} \qquad (2.2\text{-}38)$$

Critical values for $T_{n(1-\alpha)}$ are reported in Table A-10. For $n > 25$ we can use the normal approximation:

$$T_{n(1-\alpha)} = \frac{n(n+1)}{4} - z_{(1-\alpha)}\sqrt{\frac{1}{24}n(n+1)(2n+1)} \quad . \qquad (2.2\text{-}39)$$

If the values of both samples appear in dichotomized form A,B (nominal scale), the n linked observations can be classified into a 4-fold table:

sample 2 \ sample 1	A	B
A	n_{AA}	n_{AB}
B	n_{BA}	n_{BB}

(2.2-40)

Put $n_{BA} + n_{AB} = n_d$ and n_{min} the smaller between n_{AB} and n_{BA}. The hypothesis,

$$H_0 : \pi_{AB} = \pi_{BA} \qquad (2.2\text{-}41)$$

i.e. the comparison of two frequencies if <u>no margin</u> is fixed (only n_d is fixed), will be rejected if:

$$n_{min} \leq L_{(\alpha)}, \qquad (2.2\text{-}42)$$

$L_{(\alpha)}$ obtained from the binomial distribution (Table A-22)*. For large values of n_d(>20), the normal approximation can be used:

$$L_{(\alpha)} = \frac{1}{2}\left(n_d + z_{(\frac{\alpha}{2})}\sqrt{n_d}\right) \qquad (2.2\text{-}43)$$

or easier, the test of <u>Mc Nemar</u> (Conover, 1971) for significance of changes:

$$\chi^2 = \frac{(n_{AB} - n_{BA})^2}{n_d}. \qquad (2.2\text{-}44)$$

For values of $\chi^2_1 \geq \chi^2_{1(1-\alpha)}$, the null hypothesis must be rejected.

If A and B are expressed at least in ordinal scale, the above test is known as the <u>sign test</u>, even if this in its original form considered only the case with $n_{AA} = n_{BB} = 0$.

* The fact that tests (2.2-42) and (2.2-43) are based only on the non-matching responses is in some way misleading. In fact the probabilities of hypothesis $H_0 : \pi_{AB} = \pi_{BA}$ are estimated by:

$$P_{AB} = \frac{n_{AB}}{n_d} \quad \text{and} \quad P_{BA} = \frac{n_{BA}}{n_d}$$

but could also be estimated by:

$$P'_{AB} = \frac{n_{AB}}{n} \quad \text{and} \quad P'_{AB} = \frac{n_{BA}}{n}$$

leading to quite different results. A simple way, though approximated, to consider the difference between n_d and n, is to substitute in (2.2-42) n_{min} by:

$$n'_{min} = \frac{n_d}{2} - \left(\frac{n_d}{2} - n_{min}\right)\sqrt{\frac{n_d}{n}} \quad . \qquad (2.2\text{-}42')$$

2.2.8 A combination of type I errors

Suppose we have applied some different and independent tests to a situation and obtain different responses for which we evaluate the type I error (P_i). Fisher (1954) showed that it is possible to estimate the significance of the type I errors, by computing:

$$\chi^2_{2m} = -2 \sum_{i=1}^{m} \ln P_i \,. \qquad (2.2\text{-}45)$$

If $\chi^2_{2m} < \chi^2_{2m(1-\alpha)}$ then the "mean" type I error of the different tests is larger than α, otherwise it is smaller.

2.3 Hypotheses about location measures

2.3.1 Test whether the mean μ, of an approximately normal population with unknown variance, equals a preassigned value μ_0

To test the hypothesis whether a mean equals a preassigned value, or equivalently, whether a sample belongs to a population with that preassigned mean, i.e.

$$H_0 : \mu = \mu_0 \qquad \text{against} \qquad H_1 : \mu \neq \mu_0 \qquad (2.3\text{-}1)$$

we need a sample of size n and evaluate its mean $\bar{x} = \frac{1}{n} \Sigma x_i$ and variance

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum x_i^2 - n \bar{x}^2 \right) .$$

Hypothesis (2.3-1) will be rejected if Student's test

$$t_{n-1} = \frac{|x - \mu_0|}{s} \sqrt{n} \geq t_{n-1(1-\frac{\alpha}{2})} \,. \qquad (2.3\text{-}2)$$

For small samples, $n < 20$, the sample standard deviation s can be approximated by the sample range R (i.e. the maximum observed minus the minimum observed value: $R = x_{(n)} - x_{(1)}$). In this case, test (2.3-2) reduces to:

$$\lambda_n = |\bar{x} - \mu_0| \,/\, R \qquad (2.3\text{-}3)$$

with $\lambda_{n(1-\alpha)}$ in Table A-11.

If in (2.3-3) the mean is estimated by the median $\tilde{x}$ of the sample, we have the test:

$$\lambda_n = \frac{|\tilde{x} - \mu_0|}{c_n R} \qquad (2.3\text{-}4)$$

where c_n is a correction factor reported in Table A-2.

If the hypothesis of normality cannot be sustained, hypothesis (2.3-1) can be tested by the analogous based on the median:

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 \qquad (2.3\text{-}5)$$

The observed values must be ordered in a ranked sequence $x_{(j)} \leq x_{(j+1)}$ and hence the above hypothesis can be tested establishing the 95% confidence interval with lower limit

$$x_{(h)} \qquad (2.3\text{-}6)$$

and upper limit

$$x_{(n-h+1)} \quad . \qquad (2.3\text{-}7)$$

Values for h at the 95% bilateral confidence level are given in Table A-2 (Sachs, 1970). If $\tilde{\mu}_0$ falls outside this interval (L,U) the null hypothesis must be rejected. This test is meaningless for $n \leq 8$, whereas for $8 < n \leq 11$ the confidence interval comprises the whole sample range.

2.3.2 Comparison of two means sampled from approximately normal populations with unknown variances

First we consider the case of unknown but equal variances. Given two samples of size n_1 and n_2 with means $\bar{x}_1, \bar{x}_2$ and variances s_1^2, s_2^2, we reject the hypothesis that the means do not differ significantly, or equivalently that the samples belong to the same population:

$$H_0 : \mu_1 = \mu_2 = \mu \qquad \text{against} \qquad H_1 : \mu_1 \neq \mu_2 \qquad (2.3\text{-}8)$$

if,

$$t_{n_1+n_2-2} = \frac{|\bar{x}_1 - \bar{x}_2|}{s} \sqrt{\frac{n_1 n_2}{n_1+n_2}} \geq t_{n_1+n_2-2(1-\frac{\alpha}{2})} \qquad (2.3\text{-}9)$$

where,

$$s = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2} = \frac{\sum_{i=1}^{2} \sum_{j=1}^{n_i} (x_{ij}-\bar{x}_i)^2}{n_1 + n_2 - 2} \quad .$$

If the population variances cannot be supposed to be equal, it is still possible to test exactly hypothesis (2.3-8) if both samples are of the same size ($n_1 = n_2 = r$). Then we reject the null hypothesis if the sample statistic

$$t_{r-1} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s_1^2 + s_2^2}} \sqrt{r} \geq t_{r-1(1-\frac{\alpha}{2})} \quad . \qquad (2.3\text{-}10)$$

If finally neither variances nor sample sizes are equal, the approximated test

$$t'_{n_1+n_2-2} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (2.3\text{-}11)$$

can be applied for large n. Precision of this test is increased if the DF n_1+n_2-2 are substituted by:

$$\nu = \frac{(n_1-1)\,(n_2-1)}{d^2(n_2-1) + (1-d)^2\,(n_1-1)} \quad \text{with} \quad d = \frac{s_1^2/n_1}{s_1^2/n_1 + s_2^2/n_2} \,. \qquad (2.3\text{-}12)$$

Exact critical values of (2.3-11) for $1-\alpha = 0.95$ and 0.99 are given by Aspin (1949). If $t'_\nu \geq t_{\nu,(1-\frac{\alpha}{2})}$, H_0 is rejected.

If the smallest of n_1 and n_2 is not below 5, the situation of $n_1 \neq n_2$ and $\sigma_1^2 \neq \sigma_2^2$ can be solved elegantly by the test of Scheffè discussed in Kendall (1961) reducing randomly the number of observations of the larger sample, say n_2, to that of the smaller (n_1). Then evaluate

$$w_i = x_{1i} - \sqrt{\frac{n_1}{n_2}}\, x_{2i} \qquad (2.3\text{-}13)$$

and their mean $\bar{w}$. The hypothesis (2.3-8) can then be tested by:

$$t_{n_1-1} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\Sigma(w_i-\bar{w})^2}{n_1(n_1-1)}}} \quad ; \quad n_1 < n_2 \,. \qquad (2.3\text{-}14)$$

If the observations of the two samples are obtained pair-wise, so that $n_1 = n_2 = r$, then hypothesis (2.3-8) can be tested by:

$$t_{r-1} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{1}{r-1}\sum(d_j-\bar{d})^2}}\,\sqrt{r} \qquad (2.3\text{-}15)$$

where d_j are all the r differences $x_{1j} - x_{2j}$ d_j, $(j = 1 \ldots r)$ and $\bar{d} = \frac{1}{r}\sum d_j$.

2.3.3 Comparison of the location of two populations with any distribution

Consider two samples of n_1 and n_2 observations and transform the n_1+n_2 values in a ranked sequence (Section 2.2.7). Determine the median $\tilde{x}$ of this sequence. Then build the following 4-fold table:

	Sample 1	*Sample 2*
Number of values $> \tilde{x}$	k_1	k_2
Number of values $\leq \tilde{x}$	$n_1 - k_1$	$n_1 - k_2$

(2.3-16)

Now tests (2.2-9) and (2.2-10) can be applied.

If hypothesis (2.2-7) is rejected, also the hypothesis that the medians of both populations are equal:

$$H_0 : \tilde{\mu}_1 = \tilde{\mu}_2 \qquad (2.3\text{-}17)$$

will be rejected and the alternative hypothesis $H_1 : \tilde{\mu}_1 \neq \tilde{\mu}_2$ accepted. This test is known as the median test.

Further tests which apply to hypothesis (2.3-17) are those based on the distribution functions of the populations, which in fact comprise the hypothesis of equal location measures as the Mann-Whitney-Wilcoxon (2.2-32), the run (2.2-35) and the Kolmogoroff-Smirnov test (2.2-19) with critical values (2.2-37). If the observations are obtained pair-wise, Wilcoxon's pair difference test (2.2-38) or the sign test in its different forms (2.2-42)-(2.2-44) can be used.

2.3.4 Comparison of the means of c samples

The comparison of c means, i.e. whether the means of c samples can be considered equal, is equivalent to test the null hypothesis that the samples belong to only one population, or that the means of the populations from which they derive are equal. Hence the hypothesis to be tested can be written:

$$H_0 : \mu_1 = \mu_i \text{ all } i \quad \text{against} \quad H_1 : \mu_1 \neq \mu_i \text{ some } i. \qquad (2.3\text{-}18)$$

The samples are best written in a two-way table like (2.3-19). The hypothesis (2.3-18) will be rejected if the population distribution can be assumed to be normal and with equal variances if the test (2.3-20) is larger than the corresponding tabulated value $F_{c-1,n-c(1-\alpha)}$.

Samples	1	...	i	...	c	
Observations per sample	x_{11}	...	x_{i1}	...	x_{c1}	
	$\vdots$		$\vdots$		$\vdots$	
	x_{1j}	...	x_{ij}	...	x_{cj}	
	$\vdots$		$\vdots$		$\vdots$	
	x_{1n_1}	...	x_{in_i}	...	x_{cn_c}	
Number of observations n_i	n_1	...	n_i	...	n_c	$n = \sum n_i$
Means $\bar{x}_i$	$\bar{x}_1$	...	$\bar{x}_i$	...	$\bar{x}_c$	$\bar{x} = \frac{1}{n} \sum \bar{x}_i n_i$

(2.3-19)

$$F_{c-1,n-c} = \frac{\sum(\bar{x}_i - \bar{x})^2 \, n_i}{\sum\sum (x_{ij}-\bar{x}_i)^2} \cdot \frac{n-c}{c-1} = \frac{\sum \bar{x}_i^2 \, n_i - n\bar{x}^2}{\sum\sum x_{ij}^2 - \sum \bar{x}_i^2 n_i} \cdot \frac{n-c}{c-1} . \qquad (2.3\text{-}20)$$

This type of analysis of the means is the well known 1-way analysis of variance. The study of other analysis of variance models is beyond the intention of this introductory notes on statistical tests. We refer to the already quoted books of biometrics and others about analysis of variance and covariance, like Dunn and Clark (1974).

Some models have been treated by the author (Möller, 1976) and in Sections 1.4.1 and 1.4.2.

If the observations are expressed as percentages, two ways can be followed, either transforming the frequencies as *arc sin*, as for (2.2-11), using then the ANOVA methods, or transforming them to $n_{ij} \ln n_{ij}$ as for the G-test (2.2-16) and using then the analysis of entropy methods (Garner and McGill, 1956).

If the normality assumption cannot be held, the comparison of c samples can be performed with the test of Kruskal and Wallis (1952, 1953). Therefore we transform the n observations of table (2.3-19) in a ranked sequence from 1 to n, as in Section 2.2.7, and substitute x_{ij} by their corresponding ranks R_{ij}. We form at the bottom of table (2.3-19) the sums of the ranks $R_i = \sum_j R_{ij}$. As a check we must have $\sum R_i = \frac{1}{2} n(n+1)$. Hypothesis (2.3-18) is now changed to the non-parametric hypothesis:

H_0: all c samples derive from the same population, against (2.3-21)

H_1: some of the c samples do not derive from the same population

and can be tested by:

$$\chi^2_{c-1} = \frac{12}{n(n+1)} \sum \frac{R_i^2}{n_i} - 3(n+1) . \qquad (2.3\text{-}22)$$

If $\chi^2_{c-1} \geq \chi^2_{c-1(1-\alpha)}$ the hypothesis will be rejected.

More complex models are treated by Friedman (1937, 1940), for which we report only the 2-way analysis for which no entries must be missing. In a scheme like (2.3-19) with $n_i = r$ the x_{ij} values of each row (or column) are substituted by their ranks from 1 to c, (or from 1 to r) and the sum of the ranks of each column R_i. (or row, $R_{.j}$) are calculated. For $c > 3$ (or $r > 3$) the hypothesis of no influence of the columns (or rows) can be tested by the test of Friedman:

$$\chi^2_{c-1} = \frac{12}{cr\,(c+1)} \sum_i R_i^2 - 3r\,(c+1). \qquad (2.3\text{-}23)$$

If $\chi^2_{c-1} \leq \chi^2_{c-1(1-\alpha)}$ the hypothesis can be accepted.

2.3.5 Multiple *a posteriori* comparison

It will be noted that the alternative hypothesis of (2.3-18) namely $H_i: \mu_1 \neq \mu$, has not to be valid for all i but only for some of them. That means, that only one mean can be sufficient to have the null hypothesis rejected. It would be therefore interesting to have some criteria to individuate those means, which increase mostly the numerator of F, i.e. $\Sigma(x_i - \bar{x})^2 n_i$, the sum of squares, SS, between samples. This can be achieved by the following tests. First determine the critical value $F_{c-1,n-c(1-\alpha)}$ then we get from (2.3-20) the critical value for the between SS:

$$SS_{B_{(1-\alpha)}} = \frac{\sum\sum (x_{ij} - \bar{x}_i)^2}{n - c} F_{c-1,n-c(1-\alpha)}. \qquad (2.3\text{-}24)$$

The ratio $\frac{\sum\sum (x_{ij} - x_i)^2}{n - c} = MS_w$ is the "within mean squares" (= within SS divided by its DF), so that:

$$SS_{B_{(1-\alpha)}} = (c - 1) \cdot MS_w \cdot F_{c-1,n-c(1-\alpha)}. \qquad (2.3\text{-}25)$$

Any sub-set of samples can now be tested computing simply the between SS due to that sub-set and comparing it with $SS_{B_{(1-\alpha)}}$. If it is larger than $SS_{B_{(1-\alpha)}}$ the sub-set cannot be considered to derive from the same population. This is the SS simultaneous test procedure of Gabriel (1964).

Another way of carrying out *a posteriori* tests, if all n_i = constant = r, is to use the largest difference found among a set of means (their range) in place of their SS. This sample range is then compared with the least significant range, or Tukey's multiple comparison test:

$$LSR_{(1-\alpha)} = Q_{c,n-c(1-\alpha)} \sqrt{\frac{MS_w}{r}} \qquad (2.3\text{-}26)$$

where $Q_{k,\nu}$ is the studentized range, for which critical values are reported in Table A-12.

It may be noticed that for $c = 2$,

$$Q_{2,n-2} = \sqrt{2\, t_{n-2}} \,. \qquad (2.3\text{-}27)$$

If we wish to compare a pair of means that do not have the same sample size, LSR must be modified by replacing in (2.3-26) r by $\frac{2\, n_1\, n_2}{n_1 + n_2}$, which is an average of the two sample sizes. For testing three or more means, this test is no longer exact.

In both methods, the SS simultaneous test procedure (2.3-25) and the least significant range (2.3-26), the type I error is clearly smaller than α. So these tests are on the "safe decision side". A step-wise method to bring the type I error closer to α is the Student-Newman-Keuls procedure. The sample means are written in increasing (or decreasing) order $(\bar{x}_{(1)} \leq \bar{x}_{(2)} \leq \cdots \leq \bar{x}_{(c)})$ in the head column and row of a two-fold table (2.3-28) with the number of observations. If the order is increasing, we compute the differences row mean - column mean, and report them in the lower part of the table:

means	number of observations	$\bar{x}_{(1)}$ n_1	$\bar{x}_{(2)}$ n_2	$\bar{x}_{(3)}$ n_3	...	$\bar{x}_{(c)}$ n_c
$\bar{x}_{(1)}$	n_1	0				
$\bar{x}_{(2)}$	n_2	$\bar{x}_{(2)}-\bar{x}_{(1)}$	0			
$\bar{x}_{(3)}$	n_3	$\bar{x}_{(3)}-\bar{x}_{(1)}$	$\bar{x}_{(3)}-\bar{x}_{(2)}$	0		
⋮	⋮	⋮	⋮	⋮		
$\bar{x}_{(c)}$	n_c	$\bar{x}_{(c)}-\bar{x}_{(1)}$	$\bar{x}_{(c)}-\bar{x}_{(2)}$	$\bar{x}_{(c)}-\bar{x}_{(3)}$	...	0

(2.3-28)

A difference $x_{(i)}-x_{(g)}$ in the above table to be significant must be equal to or greater than:

$$LSR_{(1-\alpha)} = Q_{k,\nu(1-\alpha)} \sqrt{MS_w} \sqrt{\frac{n_i + n_g}{2 n_i\, n_g}} \qquad (2.3\text{-}29)$$

with $k = 1 + (i-g)$ and $\nu = n-c$.

Other ways of finding classes (clusters) of similar means are obtained by dendrograms (Möller, 1973)*.

* see also Figure 8 of example 2.7.2.5.

2.4 Hypotheses about variances

The sample variance is given by $s^2 = \frac{1}{n-1}\sum(x_i-\bar{x})^2$ and can be approximated for small $n(\leq 20)$ by the sample range R (maximum observation - minimum observation): $s = R/d_n$, with d_n in Table A-2.

2.4.1 Test whether the variance σ^2 of an approximately normal population equals a preassigned value σ_0^2

The hypothesis whether the population variance σ^2 can be assumed to be σ_0^2 is equivalent to the hypothesis whether a sample with variance s^2 has been extracted from a population with variance σ_0^2. To test the null hypothesis, ignoring the population mean,

$$H_0 : \sigma^2 = \sigma_0^2 \qquad \text{against} \qquad H_1 : \sigma^2 \neq \sigma_0^2 \tag{2.4-1}$$

we evaluate,

$$\chi^2_{n-1} = \frac{(n-1)s^2}{\sigma_0^2} \quad . \tag{2.4-2}$$

The null hypothesis will be rejected if,

$$\chi^2_{n-1} \leq \chi^2_{n-1(\frac{\alpha}{2})} \qquad \text{or} \qquad \chi^2_{n-1} \geq \chi^2_{n-1(1-\frac{\alpha}{2})} \; . \tag{2.4-3}$$

The same hypothesis can be tested if s is approximated by R with the test statistic,

$$z = g_n \left| \frac{R}{d_n \sigma_0} - 1 \right| \tag{2.4-4}$$

with d_n and g_n in Table A-2.

Hypothesis H_0 is rejected if $z \geq z_{(1-\frac{\alpha}{2})}$.

2.4.2 Comparison of two variances of approximately normal populations

We will test the hypothesis whether two population variances σ_1^2 and σ_2^2 are equal, or equivalently (if the means μ_1 and μ_2 are equal) whether the two samples derive from the same population, i.e.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2 \qquad \text{against} \qquad H_1 : \sigma_1^2 \neq \sigma_2^2 \ . \tag{2.4-5}$$

The test statistic is Fisher's sample variance ratio:

$$F_{n_1-1,n_2-1} = \frac{s_1^2}{s_2^2} \qquad ; \qquad s_1^2 \geq s_2^2 \ . \tag{2.4-6}$$

If $F \geq F_{(1-\alpha)}$ the null hypothesis is rejected.

For large samples ($n_1 \geq 100$) hypothesis (2.4-5) can be tested by:

$$z = \frac{|s_1 - s_2|}{\sqrt{s_1^2 n_2 + s_2^2 n_1}} \sqrt{2n_1 n_2} \ . \tag{2.4-7}$$

If $z \geq z_{(1-\frac{\alpha}{2})}$ the null hypothesis is rejected.

2.4.3 Comparison of c variances of approximately normal populations: tests of homoscedasticity

To test the hypothesis whether c variances can be considered equal,

$$H_0 : \sigma_i^2 = \sigma^2 \qquad ; \qquad i = 1, 2, \ldots, c. \tag{2.4-8}$$

against the alternative hypothesis,

$$H_0 : \sigma_i^2 \neq \sigma^2 \qquad ; \qquad \text{some } i \tag{2.4-9}$$

we consider c samples and order them like (2.3-19), then we evaluate the sample variances s_i^2 and the pooled variance

$$s^2 = \frac{\Sigma(n_i - 1)s_i^2}{n - c} = \frac{\Sigma\Sigma(x_{ij} - \bar{x}_i)^2}{n - c} \ .$$

Hypothesis (2.4-8) can be tested by Bartlett's test (which is rather sensitive if the normality condition is not fulfilled):

$$\chi^2_{c-1} = -\frac{1}{k} \sum_{i=1}^{c} (n_i - 1) \ln \frac{s_i^2}{s^2}$$

$$= \frac{2.3026}{k} \left[(n-c) \operatorname{Log} s^2 - \sum (n_i - 1) \operatorname{Log} s_i^2 \right] \tag{2.4-10}$$

with

$$k = 1 + \frac{1}{3(c-1)} \left(\sum \frac{1}{n_i - 1} - \frac{1}{n-c} \right).$$

If $\chi^2_{c-1} \geq \chi^2_{c-1(1-\alpha)}$ H_0 is rejected.

Much easier tests, if all samples have the same size $n_i = r$ (all i) are Hartley's F_{max} test and Cochran's g-test (Cochran, 1941). To obtain F_{max} we simply divide the maximum observed variance $s^2_{(c)}$ by the minimum observed variance $s^2_{(1)}$ so as to get:

$$F_{max,c,r-1} = \frac{s^2_{(c)}}{s^2_{(1)}}. \qquad (2.4\text{-}11)$$

If $F_{max} \geq F_{max(1-\alpha)}$ (Table A-13), hypothesis (2.4-8) must be rejected.

To obtain Cochran's $g_{c,r}$-test we divide the maximum sample variance by the sum of all observed variances, hence obtaining:

$$g_{c,r} = \frac{s^2_{(c)}}{\sum_{i=1}^{c} s^2_i}. \qquad (2.4\text{-}12)$$

If $g_{c,r} \geq g_{c,r(1-\alpha)}$, nomogram by Stange and Henning (1966) reported in Figure 6, the null hypothesis must be rejected.

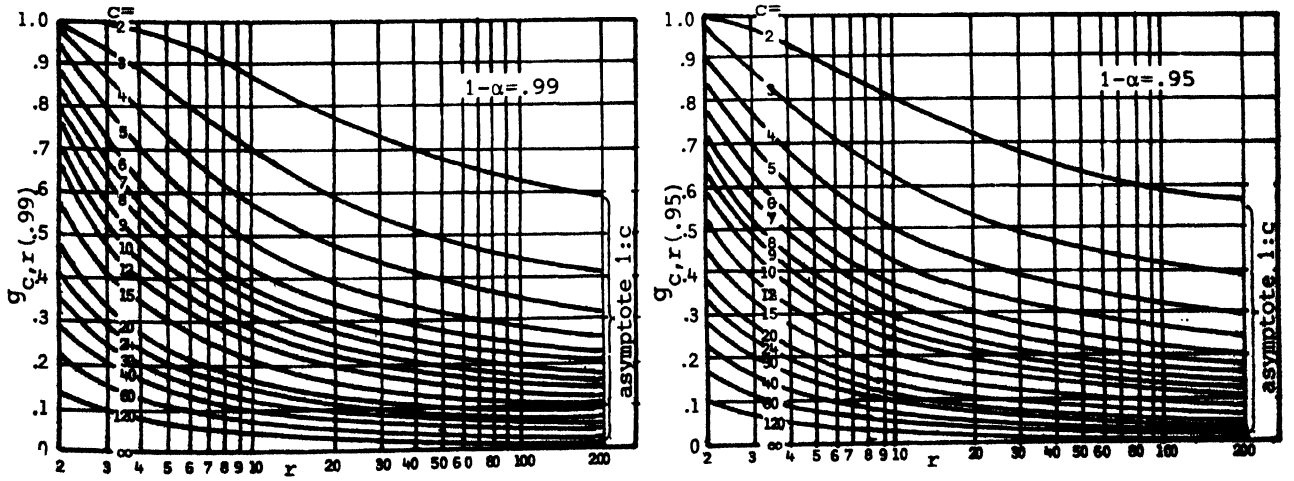


Figure 6 Critical values of Cochran's $g_{c,r(1-\alpha)}$ with confidence levels of $1-\alpha = .95$ and $1-\alpha = .99$
(from Stange and Henning, 1966, with kind permission of Springer Verlag, Berlin

2.4.4 Comparison of c variances of any distribution

A final test proposed by Scheffé (1959) consists in dividing randomly all samples in at least two sub-samples of any size (≥ 2), say in q_i sub-samples with n_{ij} observations each, $\sum_{j=1}^{q_i} n_{ij} = n_i$. For each sub-sample, evaluate the variance s_{ij}^2 and its normalizing transformation,

$$z_{ij} = \frac{n_{ij} - 1}{\sqrt{n_{ij}}} \log s_{ij}^2 . \qquad (2.4\text{-}13)$$

If $n_{ij} = m$ = constant, it is sufficient to put $z_{ij} = \log s_{ij}^2$. Now we evaluate the test statistic:

$$F_{c-1,\Sigma q_i - c} = \frac{\sum q_i (\bar{z}_i - \bar{\bar{z}})^2}{\sum\sum (z_{ij} - \bar{z}_i)^2} \cdot \frac{\sum q_i - c}{c - 1} \qquad (2.4\text{-}14)$$

which is the same test of a one-way analysis of variance (2.3-20).

If $F \geq F_{(1-\alpha)}$ hypothesis (2.4-13) must be rejected.

2.5 Hypotheses about the interdependence (correlation) and the dependence (regression) in bivariate variables

A bivariate distribution (x,y) can be characterized by five parameters μ_x, μ_y, σ_x, σ_y and the covariance σ_{xy} or the correlation coefficient

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y} .$$

In this section we describe some tests dealing with hypotheses about the linear correlation coefficient ρ. The covariance σ_{xy} is estimated through a sample of size n (= n pairs of observations x_i and y_i) by $s_{xy} = \frac{1}{n-1} \Sigma(x_i - \bar{x})(y_i - \bar{y})$ and the correlation coefficient by:

$$r = \frac{s_{xy}}{s_x s_y} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 (y_i - \bar{y})^2}}$$

2.5.1 Test of no correlation in a bivariate normal population

To test the hypothesis

$$H_0 : \rho = 0 \qquad \text{against} \qquad H_1 : \rho \neq 0 \qquad (2.5\text{-}1)$$

the test statistic is:

$$t_{n-2} = \frac{|r|}{\sqrt{1 - r^2}} \sqrt{n-2} \qquad (2.5\text{-}2)$$

If $t \geq t_{n-2,\alpha}$, H_0 is rejected.

2.5.2 Test whether the correlation coefficient ρ of a bivariate normal population equals a preassigned value ρ_0

The size of the sample to test this hypothesis must not be smaller than 25. The following functions must then be evaluated:

$$\zeta(r) = \frac{1}{2} \ln \frac{1+r}{1-r} \qquad \textit{and} \qquad \zeta(\rho_0) = \frac{1}{2} \ln \frac{1+\rho_0}{1-\rho_0} \,. \tag{2.5-3}$$

The function ζ is commonly indicated as the z-transformation of r (Fisher, 1954). The hypothesis

$$H_0 : \rho = \rho_0 \qquad \text{against} \qquad H_1 : \rho \neq \rho_0 \tag{2.5-4}$$

can then be tested by the normal variate:

$$z = \frac{1}{\sqrt{n-3}} \left[\zeta(r) - \zeta(\rho_0)\right] \tag{2.5-5}$$

If $z \geq z_{(1-\frac{\alpha}{2})}$, the null hypothesis must be rejected.

2.5.3 Comparison of two correlation coefficients of bivariate normal populations

The comparison of two correlation coefficients leads to the hypothesis:

$$H_0 : \rho_1 = \rho_2 \qquad \text{against} \qquad H_1 : \rho_1 \neq \rho_2 \,. \tag{2.5-6}$$

If the sample sizes are not smaller than 25, using transformation (2.5-3), the null hypothesis can be tested by:

$$z = \frac{\left| \zeta(r_1) - \zeta(r_2) \right|}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \tag{2.5-7}$$

where r_1 and r_2 are the correlation coefficients of two samples of size n_1 and n_2.

If $z \geq z_{(1-\frac{\alpha}{2})}$, the null hypothesis must be rejected.

If $n \leq 25$, the distribution of r can be inspected in the tables of David (1954).

2.5.4 Tests of no correlation in bivariate populations with any distribution

When the hypothesis of a normal distribution cannot be assumed, other ways of testing the independence are given by Spearman's and Kendall's rank correlation coefficients. For both the null hypothesis can be written:

H_0 : no interdependence (dependence) between x and y

against (2.5-8)

H_1 : x and y are dependent.

Given a bivariate sample (x,y) of size n, we order the x_i values in an ordered sequence of ranks R_{x_i} (from 1 to n) and so the y_i values to give R_{y_i}.

If $x_i \neq x_j$ and $y_i \neq y_j$ for all $i \neq j$, it follows that $R_{x_i} \neq R_{x_j}$ and $R_{y_i} \neq R_{y_j}$.

Spearman's rank correlation coefficient is given by:

$$r_s = 1 - \frac{6 \sum (R_{x_i} - R_{y_i})^2}{n(n^2 - 1)} \qquad (2.5\text{-}9)$$

and hypothesis (2.5-8) can be tested for $n \geq 20$ by:

$$z = |r_s| \sqrt{n-1}\,. \qquad (2.5.10)$$

For $10 \leq n \leq 20$, we have:

$$t_{n-2} = \frac{|r_s|}{\sqrt{1 - r_s^2}} \sqrt{n-2}\,. \qquad (2.5\text{-}11)$$

If $z \geq z_{(1-\frac{\alpha}{2})}$ or $t_{n-2} \geq t_{n-2,(1-\frac{\alpha}{2})}$, hypothesis (2.5-8) must be rejected.

For $4 < n < 10$, exact critical values are reported in Table A-14.

To obtain Kendall's correlation coefficient τ, we define the new variables:

$$\xi_{ij} = \begin{cases} 1 & \text{if } R_{x_j} - R_{x_i} > 0 \\ -1 & \text{if } R_{x_j} - R_{x_i} < 0 \end{cases} \quad ; \; j > i$$

and (2.5-12)

$$\eta_{ij} = \begin{cases} 1 & \text{if } R_{y_j} - R_{y_i} > 0 \\ -1 & \text{if } R_{y_j} - R_{y_i} < 0 \end{cases} \quad ; \; j > i.$$

Then Kendall's τ is defined by:

$$\tau = 1 - \frac{2 \sum_{j>i} \xi_{ij}\, \eta_{ij}}{n(n-1)} \,. \qquad (2.5\text{-}13)$$

To evaluate the sum,

$$\sum_{j>i} \xi_{ij}\, \eta_{ij}$$

it is convenient to write the ranks R_{x_i} in increasing order and the corresponding R_{y_i} ranks below them, obtaining the sequencies:

ranks of x : $1 \quad 2 \quad 3 \quad \cdots \quad R_{x_i} \quad \cdots \quad n$

ranks of y : $R_{y_{(1)}} \quad R_{y_{(2)}} \quad R_{y_{(3)}} \quad \cdots \quad R_{y_{(i)}} \quad \cdots \quad R_{y_{(n)}}$.

Now we count the number of all ranks $R_{y_{(j)}}$ on the right side of $R_{y_{(i)}}$ which are larger than $R_{y_{(i)}}$ obtaining a number which we will indicate by S_i. Then the total sum of all terms is:

$$S = \sum_{i=1}^{n-1} S_i \,, \quad (S_n = 0), \quad \text{and} \quad \sum_{j>i} \xi_{ij}\, \eta_{ij} = 2\,S - \frac{1}{2}\, n(n-1)$$

so that (2.5-13) transforms to:

$$\tau = \frac{4\,S}{n(n-1)} - 1 \,. \qquad (2.5\text{-}14)$$

Hypothesis (2.5-8) must be rejected for $n > 10$, if:

$$\frac{|\tau|}{\sqrt{2(2n+5)}} \cdot \sqrt{9n(n-1)} \geq z_{(1-\frac{\alpha}{2})} \,. \qquad (2.5\text{-}15)$$

For $n \leq 10$, exact critical values for τ are reported in Table A-14.

If the condition $x_i \neq x_j$ and/or $y_i \neq y_j$ for all $i \neq j$ is not fulfilled, but instead same x values are repeated t_x times, we associate to x_i the mean rank. Correspondingly we associate to y values repeated t_y times the mean rank. Then we form:

$$T_x = \frac{1}{2} \sum t_x(t_x^2-1) \quad \text{and} \quad T_y = \frac{1}{2} \sum t_y(t_y^2-1) \qquad (2.5\text{-}16)$$

and get Spearman's rank correlation test corrected for ties,

$$r_s = 1 - \frac{6 \sum (R_{x_i} - R_{y_i})^2}{n(n^2-1) - (T_x + T_y)} \,. \qquad (2.5\text{-}17)$$

Hypothesis (2.5-8) must be rejected for $n \geq 20$ and with the entity of ties small with respect to n, if:

$$z = |r_s| \sqrt{n-1} \geq z_{(1-\frac{\alpha}{2})} . \qquad (2.5\text{-}18)$$

In the same manner it is possible to obtain Kendall's rank correlation test corrected for ties putting $w_x = \sum_{t_x} t_x(t_x-1)$ and $w_y = \sum_{t_y} t_y(t_y-1)$:

$$\tau = \frac{2 \sum_{j>i} \xi_{ij} \eta_{ij}}{\sqrt{n(n-1)-w_x} \sqrt{n(n-1)-w_y}} \qquad (2.5\text{-}19)$$

in which ξ_{ij} and η_{ij} are defined as (2.5-12) including the zero.

Hypothesis (2.5-8) must be rejected for $n > 10$ if (2.5-15) is true.

2.5.5 Hypotheses about regression coefficients of normal bivariate populations

A regression line is defined by $y = \alpha + \beta x$. As for the correlation we can now prove easily the following hypotheses:

$$H_0 : \beta = 0 \quad \text{against} \quad H_1 : \beta \neq 0 \qquad (2.5\text{-}20)$$

$$H_0 : \beta = \beta_0 \quad \text{against} \quad H_1 : \beta \neq \beta_0 \qquad (2.5\text{-}21)$$

$$H_0 : \beta_1 = \beta_2 \quad \text{against} \quad H_1 : \beta_1 \neq \beta_2 \qquad (2.5\text{-}22)$$

First we define the following sample functions:

(i) The estimation of the regression coefficient β:

$$b = \frac{s_{xy}}{s_x^2}$$

(ii) The residual variance:

$$s_R^2 = \frac{n-1}{n-2} (s_y^2 - b^2 s_x^2) = \frac{n-1}{n-2} s_y^2 (1 - r^2) .$$

Hence we reject hypothesis (2.5-20) if the test statistic,

$$t_{n-2} = \frac{|b|}{s_R} \sqrt{(n-1)s_x^2} \geq t_{n-2,(1-\frac{\alpha}{2})} . \qquad (2.5\text{-}23)$$

Hypothesis (2.5-21) will be rejected if

$$t_{n-2} = \frac{|b-\beta_0|}{S_R}\sqrt{(n-1)s_x^2} \geq t_{n-2,(1-\frac{\alpha}{2})} \qquad (2.5\text{-}24)$$

and hypothesis (2.5-22) if

$$t_{n_1+n_2-4} = \frac{|b_1-b_2|}{S_R\sqrt{\frac{1}{(n_1-1)s_{x_1}^2}+\frac{1}{(n_2-1)s_{x_2}^2}}} \geq t_{n_1+n_2-4,(1-\frac{\alpha}{2})}. \qquad (2.5\text{-}25)$$

For further tests about regression lines we refer to Section 1.5 and Möller (1976) or any books dealing with analysis of covariance, like Searle (1971), Dunn and Clark (1974), besides the already quoted books about biometrics.

2.6 Other tests

2.6.1 Tests for outliers

Observations that differ very much from the rest of the numerical values are usually called outliers. To obtain tests to detect outliers in a sample from a normal population, the sample observations are first ordered from smallest to largest. Let $x_{(1)}$ be the smallest value and $x_{(n)}$ the largest. The hypothesis,

$$H_0: x_{(1)} \text{ is not an outlier}$$

against (2.6-1)

$$H_1: x_{(1)} \text{ is an outlier } (\equiv \text{ the distribution of the population is not normal})$$

can be tested by Dixon's test,

$$r_n = \frac{x_{(2)}-x_{(1)}}{x_{(n)}-x_{(1)}}. \qquad (2.6\text{-}2)$$

Analogously, the hypothesis

$$H_0: x_{(n)} \text{ is not an outlier}$$

against (2.6-3)

$$H_1: x_{(n)} \text{ is an outlier}$$

can be tested by:

$$r_n = \frac{x_{(n)}-x_{(n-1)}}{x_{(n)}-x_{(1)}}. \qquad (2.6\text{-}4)$$

If $r_n \geq r_{n(1-\alpha)}$, hypotheses (2.6-1) or (2.6-3) must be rejected. The critical r_n values are reported in Table A-15.

Another test for outliers based on the same assumption of normality is Grubbs' test. If the largest value is suspected we calculate the test statistic:

$$T_n = \frac{x_{(n)} - \bar{x}}{s} . \qquad (2.6\text{-}5)$$

Hypothesis H_0 (2.6-3) is rejected, i.e. $x_{(n)}$ is to be considered an outlier, if $T_n \geq T_{n(1-\alpha)}$. Critical values are given in Table A-16.

For $n \geq 20$, a suitable test is given by the ratio

$$Z_n = R / S . \qquad (2.6\text{-}6)$$

Critical values $Z_{n(\frac{\alpha}{2})}$ and $Z_{n(1-\frac{\alpha}{2})}$ which delimit the acceptance region can be found in Table A-17.

If $Z_{n(\frac{\alpha}{2})} < Z_n < Z_{n(1-\frac{\alpha}{2})}$, neither $x_{(1)}$ nor $x_{(n)}$ must be considered outliers.

If $Z_n \begin{cases} \leq Z_{n(\frac{\alpha}{2})} \\ \geq Z_{n(1-\frac{\alpha}{2})} \end{cases}$

$x_{(1)}$ must be considered an outlier if $\bar{x}-x_{(1)} \geq x_{(n)}-\bar{x}$

or

$x_{(n)}$ must be considered an outlier if $x_{(n)}-\bar{x} \geq \bar{x}-x_{(1)}$.

Once eliminated an outlier, the same procedure can be applied to the remaining data, clearly with modified confidence level.

2.6.2 Tests for randomness

Consider a sample from a normal population of n observations obtained in any order (e.g. time series). To test whether events occur in a random sequence or whether the probability of a given event is a function of the outcome of a previous event, we can use Neumann's mean square successive difference test (Neumann, 1941):

$$\eta_n^2 = \frac{\sum (x_{i+1} - x_i)^2}{\sum (x_i - \bar{x})^2} . \qquad (2.6\text{-}7)$$

The alternative hypothesis can be of two kinds:

(i) The sequence is not at random because of short-term critical influences. In this case the hypothesis of randomness is rejected if $\eta_n^2 \geq \eta_{n(1-\alpha)}^2$.

(ii) The sequence is not at random because of long-term influences (trend). In this case the hypothesis of randomness is rejected if $\eta_n^2 \leq \eta_{n(\alpha)}^2$.

The acceptance region delimited by the critical values of $\eta_{n(\alpha)}^2 - \eta_{n(1-\alpha)}^2$ are reported in Table A-18.

Another test about randomness, or more exactly about no correlation between successive terms of a time series, is the circular serial correlation coefficient of lag h:

$$R_{h,n} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(x_{i+h} - \bar{x})}{(n-1)\ s^2} \qquad (2.6\text{-}8)$$

where, $x_{i+h} = x_{i+h-n}$ for all terms x_{i+h} for which $i+h > n$.

The hypothesis of no serial correlation can then be accepted at the confidence level of $1-2\alpha$, if $R_{h,n}$ lies within the acceptance region $R_{h,n(\alpha)} - R_{h,n(1-\alpha)}$ as reported in Table A-18.

The hypothesis of randomness can also be tested by the test for runs up and down (Wald and Wolfowitz, 1940). Herefore the signs of all n-1 differences $x_i - x_{i-1}$ are recorded so as to give a sequence of only + and -. A sequence of the same sign is defined a run of length equal to the number of equal signs. The original sequence of x_i values is thus transformed into a sequence of r runs. The hypothesis of randomness can be tested by the normal approximation which is satisfactory for n as low as 10:

$$z = \frac{r - (2n-1)/3}{\sqrt{(16n-29)/90}} . \qquad (2.6\text{-}9)$$

The null hypothesis of randomness can be accepted if

$$z_{(\frac{\alpha}{2})} < z < z_{(1-\frac{\alpha}{2})} .$$

For a sequence of dichotomized data a test of randomness is the run test (2.2-35).

To test a time series specifically against the presence of a trend, the sign test (2.2-42 to 44), known under this circumstance as the Cox and Stuart (1955) test for trend, is applicable. Just group the variables into pairs:

$$(x_i \;;\; x_{i+c}) \qquad i = \begin{cases} 1 \cdots c & ;\; c = \frac{n}{2} \;;\; n \text{ even,} \\ \vdots & \\ (c-1) & ;\; c = \frac{n+1}{2} \;;\; n \text{ odd .} \end{cases}$$

Then evaluate all c or c-1 differences $x_i - x_{i+c}$ and apply the sign tests (2.2-42) or (2.2-43).

2.7 Examples

2.7.1 Examples of tests about frequencies

2.7.1.1 Examples of tests about single frequencies

Consider the first sample of *Mullus barbatus* of Table V. Is it possible to accept the hypothesis that 50% of the fish have a concentration level of .1($\equiv H_0$) or are there significantly more ?

By inspection of the table we get:

$$\begin{aligned} k &= 7 \quad \text{fish with a concentration higher than .1} \\ \underline{n-k} &= 3 \quad \text{fish with a concentration lower than .1} \\ n &= 10 \end{aligned}$$

The hypothesis to be tested is $H_0: \pi = .5$ against $H_1: \pi > .5$, and the appropriate test (2.2-2):

$$F_{8,14} = \frac{.5}{1-.5}\,\frac{7}{4} = 1.75 \leq F_{8,14(.95)} = 2.70 .$$

Hence, H_0 can be accepted: The difference between the observed frequency of $\frac{7}{10} = .7$ and the hypothesized value of $\pi = .5$ cannot be considered to be high enough to state that significantly more than 50% of the fish have a concentration level of .1.

Since $\pi_0 = .5 \leq \frac{12}{22}$ the above test cannot be approximated by (2.2-3).

If we consider only the first concentration levels of all fishes, we get:

$$\begin{aligned} k &= 44 \\ \underline{n-k} &= 37 \\ n &= 81 . \end{aligned}$$

Since

$$\frac{12}{12+n} - \frac{12}{93} = 0.13 < \pi = .5$$

now test (2.2-3) can be used, evaluating

$$p = \frac{44}{81} = .5432$$

and hence

$$z = \frac{.5432 - \frac{1}{162} - .5}{\sqrt{.5 \cdot .5}} \sqrt{81} = .67 < z_{(.95)} = 1.645 .$$

Again, H_0 can be accepted: As in the previous case, we can conclude that the difference between $p = .5432$ and $\pi = .5$ cannot be considered to be high enough to state that significantly more than 50% of the fish have a concentration level of .1.

2.7.1.2 Example of a comparison of two frequencies by Fisher's exact test

We compare the frequencies as in 2.7.1.1 of samples 1 and 2 of Table V. We have:

$k_1 = 7$	$k_2 = 9$	$k = 16$
$n_1 - k_1 = 3$	$n_2 - k_2 = 1$	$n - k = 4$
$n_1 = 10$	$n_2 = 10$	$n = 20$

Since n is not too large and in the 4-fold table figures one very small entry, we use the exact test of Fisher (2.2-9) evaluating:

$$P_1 = \frac{16!\ 4!\ 10!\ 10!}{7!\ 9!\ 3!\ 1!\ 20!} \quad ; \quad P_0 = \frac{16!\ 4!\ 10!\ 10!}{6!\ 10!\ 4!\ 0!\ 20!} .$$

Recurring to Table A-1 we get the logarithms:

$$\begin{aligned} \text{Log } P_1 &= (13.3206 + 1.3802 + 2 \cdot 6.5598) - (3.7024 + 5.5598 + .7782 + 18.3861) \\ &= 27.8204 - 28.4265 = -.6061 \\ P_1 &= .248 \\ \text{Log } P_0 &= 27.8204 - (2.8573 + 6.5598 + 1.3802 + 18.3861) \\ &= -1.363 \\ P_0 &= .043 \end{aligned}$$

Hence $P = P_1 + P_0 = .248 + .043 = .291$, which, being larger than $\frac{\alpha}{2} = .025$ permits to accept the hypothesis (2.2-7):

$$H_0 : \pi_1 = \pi_2 \qquad \text{against} \qquad H_1 : \pi_1 \neq \pi_2 .$$

That means that the observed frequencies of $\frac{7}{10}$ for the first and $\frac{9}{10}$ for the second sample do not differ significantly.

2.7.1.3 Example of a comparison of two frequencies by χ^2 test and arc sin transformation

Bulnheim (1977) compared the survival rates of *Gammarus duebeni* descending from inbred and outbred strains. From a number of 47 outbreeding pairs he obtained $n_1 = 6809$ young and $k_1 = 6065$ survived up to the stage of sub-adults. From a number of 33 inbreeding pairs $n_2 = 2638$ young were obtained, of which $k_2 = 2013$ reached the stage of sub-adulthood. Can the survival rates,

$$p_1 = \frac{k_1}{n_1} = .89 \qquad \text{and} \qquad p_2 = \frac{k_2}{n_2} = .69$$

be considered equal? That is the null hypothesis to be tested. Hence by (2.2-10), we obtain:

$$\chi_1^2 = \frac{[6065 \cdot 625 - 2013 \cdot 744]^2 \cdot 9447}{6809 \cdot 2638 \cdot 8078 \cdot 1369}$$

$$= 250.04 \gg \chi^2_{1(.995)} = 7.879 .$$

Hence the hypothesis of equal survival rates will be rejected.

We could also have used (2.2-11):

$$z = \frac{|\text{arc sin } \sqrt{.89} - \text{arc sin } \sqrt{.69}|}{\frac{1}{2}\sqrt{\frac{9447}{6809 \cdot 2638}}} = \frac{|1.233 - .980|}{\frac{1}{2}\sqrt{.000526}} = 22$$

which confirms the above result.

Finally we evaluate (2.2-16) first building the table of n times the normal logarithm of the entries of the four-fold table of the observed frequencies:

observed frequencies			n ln n		
6065	2013	8078	52828	15314	72677
744	625	1369	4920	4024	9887
6809	2638	9447	60096	20782	86473

Hence $G_1 = 2(52\,828 + 4\,920 + 15\,314 + 4\,024 - 72\,677 - 9\,887 - 60\,096 - 20\,782 + 86\,473) = 234$ which approximates fairly well the above χ_1^2 .

2.7.1.4 Example of a comparison of four frequency distributions by χ^2 and G-test

Temperature and salinity choice can be altered following or/and during pesticide exposure (Bayne *et al.*, 1973). The following table gives an example for salinity choice by an euryhaline organism:

Salinities chosen	Concentration of poison: 0	1.0 µg	10 µg	100 µg	
0 ‰	60	2	8	10	80
15 ‰	10	28	5	13	56
30 ‰	15	35	39	43	132
	85	65	52	66	268

To test whether with increasing pesticide exposures higher salinities are preferred or not, we can recur to (2.2-14) obtaining $\chi^2_6 = 121.1$ or (2.2-16): $G_6 = 122.06$. Both values are much higher than the critical value of $\chi^2_{6(.95)} = 12.592$.

The advantage of the G-test lies in the fact that sub-hypotheses can be easily tested. First, let us consider the logarithmic transformation of the above table which led us to G_6:

245.66	1.39	16.64	23.03	350.56
23.03	93.3	8.05	33.34	225.42
40.62	124.44	142.88	161.73	644.53
377.63	271.34	205.46	276.52	1498.38

If we were interested to check if there are any significant differences between the effect due to the three different concentrations of poison, hence eliminating the first column, all we have to do is to evaluate the new marginals of the last column which then becomes:

frequencies	$n_i \log n_i$
20	59.91
46	176.12
117	557.17
183	953.34

Now $G_4 = 23.2$ which is still above the 5% critical value denoting that the three chosen concentration values determine a different response. With the same method, we obtain instead for the concentration of 10 µg and 100 µg, $G_2 = 2.44$, which is smaller than the critical value: there is no significant difference in the responses.

2.7.1.5 Example of a goodness-of-fit test

It is a common experience of field biologists that relatively low or medium colonisation densities of a given species are much more frequently encountered than high and extremely high ones. Hence the assumption of a Poisson distribution seems reasonable.

A large random sample was taken (Elliot, 1971) of the number $x_i = K$ of maifly nymphs on the tops of stones at night, obtaining the frequencies n_K. The mean is $\bar{x} = 10.125$ and the expected frequencies of the Poisson distribution could thus be determined:

$$\pi_K = \exp(-\bar{x}) \frac{\bar{x}^K}{K!} .$$

To use the χ^2 test (2.2-18) we first group the extreme classes together to satisfy the condition, so as to have as few as possible classes with expected frequencies lower than 5. The results are indicated in Table XV, from which $\chi^2 = .62$ is obtained and which, being smaller than 1, needs no inspection of any table because the correspondence between observed and expected frequencies is almost perfect. The DF of χ^2 must be lowered by 1 since one parameter, $\bar{x}$, had to be estimated (DF = r-1-1 = 9).

Hence, the colonization densities of maifly nymphs can be assumed to be distributed like the Poisson density function. Since the fundamental characteristic of the Poisson distribution is to let unchanged $\bar{x} = n\pi$, a very low colonization probability

$$\pi = \frac{\bar{x}}{n} = \frac{10.125}{80} = .127$$

follows from the observed data, confirming the idea stated at the beginning of relative low or medium colonization density. Further, it follows that, for increasing n, the probability π tends to diminish.

In the last three columns of Table XV we have reported the cumulated relative observed, $F(x_i)$, and expected, $\Phi(x_i)$, frequencies, and the absolute value of their differences $D(x_i)$ to apply the Kolmogoroff-Smirnov test (2.2-19). Hence, $D_{80} = .0224$ which is well below the critical value of $D_{80(.96)} = .167$ confirming the above results.

Table XV

Comparison of an observed with a Poisson distribution

$x_i = K$	n_K	$n\,\pi_K$	$n_K - n\,\pi_K$	χ^2	$F(x_i)$	$\Phi(x_i)$	$\lvert D(x_i)\rvert$
≤ 4	2 } 4	2.15 } 4.99	-.99	.20	.0250	.0269	.0019
5	2 }	2.84 }			.0500	.0625	.0125
6	4	4.79	-.79	.13	.1000	.1224	.0224
7	7	6.93	.07	.00	.1875	.2092	.0217
8	10	8.77	1.23	.17	.3125	.3189	.0064
9	10	9.86	.14	.00	.4375	.4424	.0049
10	10	9.99	.01	.00	.5625	.5674	.0049
11	10	9.19	.81	.07	.6875	.6824	.0051
12	8	7.75	.25	.01	.7875	.7794	.0081
13	6	6.04	-.04	.00	.8625	.8551	.0074
14	4	4.37	-.37	.03	.9125	.9098	.0027
15	4 }	2.84 }			.9625	.9466	.0159
16	2 } 7	1.87 } 7.20	.20	.01	.9875	.9700	.0175
17	1 }	1.11 }			1	.9839	.0161
≥ 18	0 }	1.28 }			1	1	0
	80	79.89		$\chi^2_9 = .62$			

2.7.1.6 Example of a probit transformation to determine LC_{50}

To determine the 96 hs LC_{50} of Syntopon C for ***Scolelepis fuliginosa***, Stora (1972) fitted by probit transformation a normal distribution to the mortality curve thus obtaining the equation of the expected probit line $y(x) = a + bx$. Solving this for $y(x) = 5$, the LC_{50} is easily obtained. The result of Stora's analysis is reported in Table XVI. Empirical probits and working probits have been obtained by (2.2-22) recurring, as for the weights, to Tables A-19 and A-20.

Table XVI

Observed mortality curve and its probit transformation
[from Stora (1972), with kind permission of Station marine d'Endoume, Marseille]

Concentration mg/l	Number of dead	Percent	x_j	$Y(x_j)$	$Y'(x_j)$	$y(x_j)$	w_j	Products			
			Logarithms of concentration	Empirical probits	Expected probits	Working probits	Weights	$w_j x_j$	$w_j y_j$	$w_j x_j y_j$	$w_j x_j^2$
9	19	95	0.95424	6.6449	6.48	6.6250	5.5120	5.2597	36.5170	34.8455	5.0190
8	17	85	0.90309	6.0364	6.08	6.0349	8.2280	7.4306	49.6551	44.8429	6.7105
7	13	65	0.84510	5.3853	5.63	5.3671	11.0000	9.2961	59.0381	49.8930	7.8561
6	11	55	0.77815	5.1257	5.10	5.1255	12.6860	9.8716	65.0220	50.5968	7.6815
5	9	45	0.69897	4.8743	4.48	4.9030	11.5260	8.0563	56.5119	39.5000	5.6311
4	2	10	0.60206	3.7184	3.73	3.7195	6.9240	4.1686	25.7538	15.5051	2.5097
						Sums:	55.8760	44.0829	292.4979	235.1833	35.4079

$$\bar{x} = \frac{44.0829}{55.876} = .789 \qquad S_x^2 = \frac{35.4079}{55.876} - .789^2 = .011166$$

$$\bar{y} = \frac{292.4979}{55.876} = 5.235 \qquad S_{xy} = \frac{235.1833}{55.876} - .789 \cdot 5.235 = .0786$$

Hence, $b = \frac{.0786}{.011166} = 7.04$

and, $a = 5.235 - 7.04 \cdot .789 = -.32$,

so that the expected probit line is defined by $y_j = -.32 + 7.04\, X_j$.

For $y_j = 5$ we obtain $x_j = \log LC_{50} = .75568$ which corresponds to a concentration of 5.7 mg/l.

2.7.1.7 Example of normality tests

We consider sample 9 of Table V and specifically the n=21 differences between determinations, obtaining (multiplying by 1000):

2, 0, 0, -3, -2, 0, 6, -7, 0, -7, -4, -1, 0, -4, -2, -1, -9, -5, -4, -1, -4.

We evaluate the skewness b_1, the kurtosis b_2, the probit line, W_n and r_n. The main operations are reported in Table XVII.

Observations:

(i) From the first 4 central moments of the x_i values, we have obtained from (2.2-25) and (2.2-26) $b_1 = .128$ and $b_2 = 3.357$. Both values lie within the acceptance area of Figures 4 and 5 for $n = 30$, hence all the more for $n = 21$. This means that we can accept the normality hypothesis. $b_1 > 0$ tells us that the frequency distribution has a longer tail to the right of the central maximum, and $b_2 > 3$ that the curve is leptokurtic.

(ii) To evaluate the probits, one has to first compute the cumulative frequencies by:

$$\frac{i - .5}{21} = \frac{21 - 1}{42}$$

with $i = 1 \cdots 21$. From Table A-19 the probits $Y(x_i)$ are obtained and reported in Figure 7. It is seen that a straight line fits rather well to the data though it seems that a sigmoid curve fits better, denoting a slight over-dispersion as to the normal (confirming b_2), which is more pronounced at the right side (confirming b_1).

(iii) For the W_n-test (2.2-28) the difference between the largest observed value, $x_{(n)}$, minus the smallest $x_{(1)}$ is formed, then after elimination of these values, the difference between the next largest $x_{(n-1)}$ and the next smallest $x_{(2)}$ is formed, and so on. Correspondingly we obtain the weights a_i from Table A-23. The sum of the products is obtained, $b = 14.7399$, and divided by $(n-1)s^2$. The result is $W_{21} = .9561$. From Table A-24 it is seen that the observed value lies well above the critical value of $W_{21(.05)}$, hence permitting to accept the hypothesis of normality.

(iv) In the last two columns of Table XVII are reported the values for determining r_n (2.2-30). First determine m_i using (2.2-29). Here it is only necessary to evaluate the upper half of the table, since the corresponding normal order statistic medians are equal to the lower part, changing the sign. The M_i values corresponding to m_i are obtained from Table A-3; then the ordinary correlation coefficient between $x_{(i)}$ and M_i is evaluated, obtaining $r_{21} = .973$ which, confirming the former results, lies well above the critical value $r_{21(.05)} = .952$, obtained from Table A-25.

Table XVII

Construction of the probit distribution and W_n- and r_n-tests for normality

Order	Ordered sequence	Probit analysis		W_n-normality test		r_n-normality test	
		Cumulated frequence	Probits	Differences	Weights	Uniform order statistic medians	Normal order statistic medians
i	$x_{(i)}$	$\frac{i-\frac{1}{2}}{21}$	$Y(x_{(i)})$	$x_{(n-i+1)}-x_{(i)}$	a_i	m_i	M_i
1	-9	.024	2.96	15	.4643	.033	-1.838
2	-7	.071	3.52	9	.3185	.079	-1.412
3	-7	.119	3.81	7	.2578	.126	-1.146
4	-5	.167	4.04	5	.2119	.172	- .927
5	-4	.214	4.21	4	.1736	.219	- .776
6	-4	.262	4.37	4	.1399	.266	- .625
7	-4	.309	4.50	4	.1092	.313	- .487
8	-4	.357	4.63	3	.0804	.360	- .343
9	-3	.405	4.76	2	.0530	.408	- .238
10	-2	.452	4.87	1	.0263	.453	- .118
11	-2	.500	5.00	-	0	.500	0
12	-1	.548	5.13				.118
13	-1	.595	5.24				.238
14	-1	.643	5.37				.343
15	0	.690	5.50				.487
16	0	.738	5.63				.625
17	0	.786	5.79				.776
18	0	.833	5.96				.927
19	0	.880	6.19				1.146
20	2	.929	6.48				1.412
21	6	.976	7.04				1.838
$\bar{x}$	-2.19			$b = 14.7399$		$r_{21} = .973$	
s_x^2	11.36						
b_1	.128			$W_{21} = \frac{14.7399}{20 \times 11.36} = .9561$			
b_2	3.357						

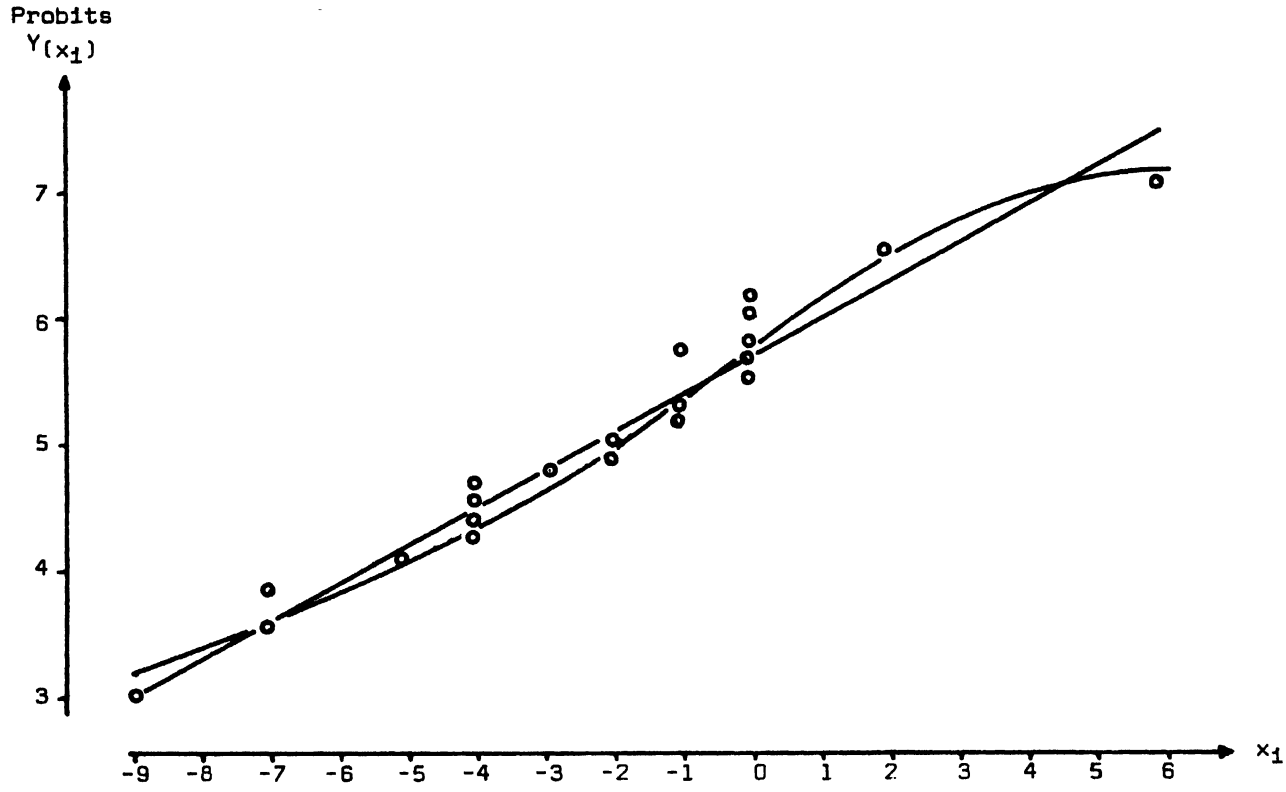


Figure 7 Probit analysis of data of Table XVII

2.7.1.8 Example of a comparison of two distributions by the Mann-Whitney U-test and van der Waerden's X-test

Hatfield and Johansen (1972) tested the learning ability of insecticide treated salmon in a shuttle box conditioning apparatus. The fish had to learn to escape from a dark chamber through a door in a lighted one. Light therefore was the conditioned, electric shocks the unconditioned stimuli. The learning ability was measured by the number of trials. Prior to the learning procedure the fish underwent a 24-hour exposure to DDT at concentrations near 96-hour LC_{50} obtaining results (Table XVIII) of which also ranks are reported to apply the Mann-Whitney-Wilcoxon U-test (2.2-32).

Table XVIII

Transformation of observed values into ranks

Control		DDT 70 µg / l	
No.of trials x_{1j}	Ranks	No.of trials x_{2j}	Ranks
7	3.5	19	15.5
14	12	15	14
3	2	14	12
9	10	8	7
8	7	19	15.5
8	7	> 75	17.5
2	1	> 75	17.5
7	3.5	8	7
		14	12
		8	7
$n_1 = 8$	$R_1 = 46$	$n_2 = 10$	$R_2 = 125$

check: $\frac{1+n}{2} \cdot n = R_1 + R_2 = 171$.

$$U_1 = 8 \cdot 10 + \frac{8 \cdot 9}{2} - 46 = 70 \quad ; \quad U_2 = 10$$

Since U_2 (the smaller between U_1 and U_2) is smaller than the critical values reported in Table A-8, we must conclude that DDT has a significant influence on the learning process of salmons.

Since the difference $|n_1 - n_2| = 2 < 5$, the same hypothesis could have been tested by van der Waerden's X-test (2.2.30), transforming the ranks of one sample into fractions

$$\frac{r_j}{n+1}$$

and obtaining then the normal variates from Table A-3. We thus get for the smaller of both samples (the first):

$\frac{r_j}{n+1}$	.18	.63	.11	.53	.37	.37	.05	.18	Sum
$Y^{-1}(x_j)$	-.915	.332	-1.227	.075	-.332	-.332	-1.645	-.915	X = 5.773

which is larger than all critical values reported in Table A-21 for n = 18, hence confirming the above results.

2.7.1.9 Example of a comparison of two distributions by the Kolmogoroff-Smirnov test

Another example of a comparison between two distributions is obtained recurring to the Kolmogoroff-Smirnov test (2.2-19) with data in accordance to Hoppenheit (1977) regarding the cumulative relative frequencies of the production of eggs per egg-sac in population weekly exploited at rates of 10% and 90%. The observed values of samples with $n_1 = 100$ and $n_2 = 200$ and the computations of the differences between the frequencies are reported in Table XIX.

Table XIX

Differences between two distribution functions

Eggs/sac	High density cumulative frequencies F_1	Low density cumulative frequencies F_2	Differences $\lvert F_1 - F_2 \rvert$
0-10	0	0	0
10-20	.05	.01	.04
20-30	.15	.03	.12
30-40	.50	.08	.42
40-50	.80	.20	.60
50-60	.90	.40	.50
60-70	1.	.70	.30
70-80	1.	.82	.18
80-90	1.	.90	.10
90-100	1.	1.	0

Comparing the largest difference with the critical value (2.2-37) we obtain

$$D_{n_1,n_2} = .60 > 1.48 \sqrt{\frac{300}{20000}} = .18$$

Hence both distributions are significantly different.

2.7.1.10 Example of a comparison of two linked distributions

In Table V are reported two determinations for each pretreatment. Hence both observations for each individual are linked and the Wilcoxon pair difference test (2.2-38) is applied.

Let us consider sample 9 and evaluate the difference between the determinations, obtaining d_j, like in example 2.7.1.7. Transform the absolute values of the non-null differences into a ranked sequence and obtain the sums of the ranks of the positive, T^+, and negative, T^-, differences. The results are summarized in Table XX:

$$(T^+ + T^- = \frac{n_d(n_d+1)}{2} = 136)$$

Table XX

Ranks of observed values

Non-null differences d_j	Ranks of absolute values
.002	5
-.003	7
-.002	5
.006	13
-.007	14.5
-.007	14.5
-.004	9.5
-.001	2
-.004	9.5
-.002	5
-.001	2
-.009	16
-.005	12
-.004	9.5
-.001	2
-.004	9.5
$n_d = 16$	18 + 118 = 136

$$T^+ = 18 \; ; \; T^- = \frac{16 \cdot 17}{2} - 18 = 118.$$

The critical value for $n_d = 16$ from Table A-10 results to be $T_{16(.975)} = 30$ hence larger than the inferior of T^+ and T^-, so that the hypothesis that both determinations lead to the same result must be rejected.

Applying the sign test (2.2-42) we have $n_{min} = 2$ of $n_d = 16$. From Table A-22 it is seen to coincide with the 1% critical value. The alternative test (2.2-42') with

$$n'_{min} = 8-(8-2)\sqrt{\frac{16}{21}} = 2.76$$

lies instead nearer to the 5% critical value, still confirming the previous results of a significant difference between the replications of the determinations.

2.7.2 Examples of tests about location measures

2.7.2.1 Examples of a comparison of observed location measures with theoretic values

The hypotheses of Chapter 2.7.1.1 that 50% of the fish have a concentration level of 0.1 can be set parametrically, like (2.3-1):

$$H_0 : \mu = .1 \quad ; \quad H_1 : \mu > .1$$

which means that the mean value of the concentrations is either 0.1 or higher.

The possible tests can be based

(i) on the sample mean $\bar{x} = .2062$ and the standard deviation $s = .2196$; (2.3-2) and Table A-4

$$t_9 = \frac{.2062-.1}{.2196}\sqrt{10} = 1.53 \leq t_{9(.95)} = 1.833$$

(ii) on the sample mean, and the sample range $R = .766-.037 = .729$; (2.3-3) and Table A-11

$$\lambda_{10} = \frac{.2062-.1}{.729} = .146 \leq \lambda_{10(.95)} = .186$$

(iii) not parametrically on the sample median $\tilde{x} = \frac{1}{2}(.135 + .167) = .151$ and the sample range; (2.3-4) and Tables A-2 and A-11

$$\lambda_{10} = \frac{.151-.1}{1.175 \cdot .729} = .060 \leq \lambda_{10(.95)} = .186$$

(iv) distribution free on the sample median and its confidence interval (L, U) defined by (2.3-6) and (2.3-7)

$$L = x_{(1)} = .037 \quad ; \quad U = x_{(n)} = .766 \ .$$

In all four cases hypothesis H_0 that the mean value of the concentration is 0.1 can be accepted: In the last case, because of the low number of observations, the confidence interval equals the sample range. To consider a less extreme case, we evaluate now the 5% confidence interval for the median of the differences of example 2.7.1.7 using (2.3-6), (2.3-7) and recurring to Table A-2, obtaining:

$$L = x_{(5)} = -4 < \tilde{\mu} < U = n_{(17)} = 0 \ .$$

The hypothesis (2.3-5) $H_0 : \tilde{\mu} = 0$, is hence acceptable as a limit situation.

Thus, sample 9 can be said to belong to a population with the preassigned concentration mean of 0.1.

2.7.2.2 Examples of a comparison of two samples

Labat et al. (1974) observed disappearance of mucus cells in gill filaments following to exposure to copper sulfate obtaining the results of Table XXI. To test whether the exposure to copper sulfate has a significant influence we can choose one of the tests of 2.3.2.

Table XXI

Visible mucus cells/100 µ

Control x_{1j}	1 mg Cu/l • 24 h x_{2j}
16	10
17	8
12	10
18	12
11	13
18	14
12	6
15	5
16	7
14	5
18	10
12	11
	9
	8
$n_1 = 12$	$n_2 = 14$
$\bar{x}_1 = 14.917$	$\bar{x}_2 = 9.143$
$s_1^2 = 6.992$	$s_2^2 = 7.978$

In order to know whether (2.3-8) or (2.3-10) is preferable we first test whether the variances can be assumed to be equal (homoscedastic), supposing normality. This is done by test (2.4-6):

$$F_{13,11}$$

Hence the variances can be assumed to be homoscedastic. Test (2.3-8) gives,

$$t_{24} = \frac{|14.917 - 9.143|}{\sqrt{\frac{11 \cdot 6.992 + 13 \cdot 7.978}{24}}} \sqrt{\frac{12 \cdot 14}{12 + 14}} = 5.35$$

Table XXII

The population density per square centimetre of Simuliidae pupae on the 6 surfaces of 30 boulders from 3 stations in the Endrick system (From Maitland and Penney, 1967, with kind permission of Blackwell Scientific Publications Ltd., Oxford)

Boulder		Front (i)	Upper (ii)	Back (iii)	Under (iv)	Side (v)	Side (vi)
A	1	Absent	0.37	0.27	0.00	0.02	0.01
	2	0.01	0.04	0.79	0.16	0.17	0.04
	3	0.09	0.96	2.52	0.02	0.25	0.66
	4	0.00	0.68	2.49	0.16	0.01	0.30
	5	0.00	0.00	1.06	0.01	0.17	0.00
	6	0.00	0.01	1.97	0.00	0.00	0.00
	7	0.00	0.19	0.49	0.05	0.15	0.10
	8	0.00	0.01	0.71	0.00	1.69	0.01
	9	0.00	0.03	3.85	0.00	0.03	0.02
	10	0.00	0.08	0.09	0.00	0.00	0.11
	11	0.01	0.22	0.01	0.00	0.00	0.01
	12	0.04	0.36	0.07	0.04	0.11	1.05
	13	0.03	0.01	0.32	0.51	0.08	0.12
	14	0.00	0.02	0.06	0.00	0.00	0.00
B	1	0.00	0.03	0.05	0.00	Absent	0.05
	2	0.04	0.14	0.04	0.08	0.02	0.02
	3	0.02	0.02	0.30	0.01	0.07	0.00
	4	0.00	0.00	0.00	0.08	0.00	0.00
	5	0.00	0.08	0.07	0.00	0.00	0.00
	6	0.03	0.05	0.13	0.00	0.06	0.00
C	1	0.08	0.08	0.43	0.00	0.10	0.10
	2	0.09	0.45	0.23	0.00	0.04	0.00
	3	0.17	Absent	0.83	0.01	0.01	0.12
	4	0.00	0.13	0.06	0.00	0.02	0.02
	5	0.04	0.15	0.00	0.01	0.05	0.00
	6	0.04	0.68	0.95	0.00	Absent	0.00
	7	0.00	0.35	0.00	0.00	0.00	0.00
	8	0.00	0.11	0.06	0.00	0.03	0.07
	9	0.16	0.44	0.15	0.00	0.16	0.06
	10	0.10	0.11	0.19	0.33	0.29	0.27

From Table A-4 it is seen that this value is higher than all critical values reported there for 24 DF, hence the difference in numbers of stainable mucus cells in gill filaments of copper exposed trout and control fish is highly significant.

In another experiment Maitland and Penney (1967) calculated a figure for the population density in terms of the number of pupal cases per square centimetre of boulder surface. Results were obtained from 30 boulders: 14 from Station A (Drumtian), 6 from Station B (Finnich) and 10 from Station C (Blanc), (Table XXII).

Table XXIII

A comparison of the mean population density of Simuliidae pupae on any two surfaces, taking into account the variation among 30 boulders examined from the Endrik system (N.B. Side (vi) was not compared with the other boulder surfaces since it was found not to be statistically different from side (v))

(From Maitland and Penney, 1967, with kind permission of Blackwell Scientific Publications Ltd., Oxford)

	Areas compared	t	Degrees of freedom	α	Conclusion
A	Front (i) with back (iii)	3.36	57	0.01	Highly significant (more pupation on back)
B	Upper (ii) with back (iii)	2.30	57	0.05-0.01	Significant (more pupation on back)
C	Under (iv) with back (iii)	3.34	58	<0.01	Highly significant (more pupation on back)
D	Side (v) with back (iii)	2.60	56	0.05-0.01	Significant (more pupation on back)
E	Side (v) with side (vi)	0.03	56	0.5	No significant difference
F	Upper (ii) with under (iv)	0.32	57	0.5	No significant difference
G	Side (v) with under (iv)	0.15	56	0.5	No significant difference
H	Front (i) with under (iv)	0.45	57	0.5	No significant difference
I	Front (i) with upper (ii)	0.37	56	0.5	No significant difference
J	Side (v) with upper (ii)	0.09	57	0.5	No significant difference
K	Side (v) with front (i)	0.17	55	0.5	No significant difference

A statistical analysis was made of the results from all three stations considered together. This was done by comparing the population density on any two surfaces, taking into account the variation among the 30 boulders examined (2.3-8). From these results (Table XXIII) it can be seen that pupation is selective for the back surface and is random between all the other surfaces. In the field experiments using bricks it was also found that pupation was selective for the back surface.

The comparison of the distributions of Table XXI can also be performed by the median test (2.3-16) first building the fourfold table:

	Sample 1	Sample 2	
Number of values $> \tilde{x} = 12$	8		10
Number of values $\leq \tilde{x} = 12$		12	16
	12	14	26

Since $\frac{2 \cdot 4}{(8+1)(12+1)} < 1$ Fisher's exact test (2.2-9) can be applied leading to P = .009 confirming the result obtained by the t-test.

Other comparisons like Scheffé's t-test, the Mann-Whitney-Wilcoxon test, the run test and the Kolmogoroff-Smirnov test could also be used. Preference to one or the other method must be given according to the acceptance or not of the normality condition.

2.7.2.3 Example of a comparison of two linked samples

Consider again Chapter 2.7.1.10. To test whether the mean concentrations of both determinations differ or not significantly, we can recur to (2.3-15). The number of differences is m = 21 (and not 16, as in Table XX. since also the 5 null differences must be considered) and the means $\bar{x}_1 = .0785$, $\bar{x}_2 = .0807$. Hence,

$$t_{20} = \frac{|.0785 - .0807|\sqrt{21}}{\sqrt{\frac{.000227}{20}}} = 2.999$$

which is larger than the 5% critical value (= 2.086) so that we reject the null hypothesis as in Chapter 2.7.1.10: The mean concentration of the two determinations differs significantl

2.7.2.4 Examples of multiple distribution free comparisons

We use as example simplified data from Siebers and Bulnheim (1977). The aim of the authors has been to test whether the rate of amino-acid uptake of the *Enchytraeus albidus* is influenced by salinity. The test organisms were exposed to four concentrations (rows matched pairs) namely 20, 50, 120 and 240 µM/l of glycine. The rates of uptake (µMoles/g/h) were measured at salinities of 0 %, 10 %, 20 % and 30 % (columns), Table XXIV.

Table XXIV

Rates of uptake (µMoles/g/h)/salinities

Glycine concentration in water	0 %	10 %	20 %	30 %
20	.0020 .0069	.0310 .0304	.262 .368	.605 .706
50	.0055 .0446	.0591 .0422	.522 .517	1.334 .998
120	.0086 .0144	.0729 .0740	.566 .627	1.805 1.583
240	.0424 .0394	.0943 .1190	.895 .644	1.826 1.710

First we do not consider the different glycine concentrations and hence treat Table XXIV as a one-way classification, so that we can apply the test of Kruskal Wallis (2.3-22). We thus transform the uptake rates into ranks from 1 to 32 and sum for column totals R_i obtaining Table XXV.

Table XXV

Ranks of rates of uptake

	0 %	10 %	20 %	30 %	
	1	7	17	22	
	3	6	18	25	
	2	12	20	28	
	11	9	19	27	
	4	13	21	31	
	5	14	23	29	
	10	15	26	32	
	8	16	24	30	
Sums R_i	44	92	168	224	$528 = \frac{1+32}{2}\ 32$

Hence,

$$\chi^2_3 = \frac{12}{32 \cdot 33 \cdot 8} (44^2 + 92^2 + 168^2 + 224^2) - 3 \cdot 33 = 27.1$$

which is highly significant: The uptake rate depends on the salinity.

Now we consider also the four glycine concentration classes and apply the Friedman test (2.3-23) transforming the uptake rates of any row into ranks from 1 to 4, and sum again for columns totals, obtaining Table XXVI, from which a between columns test after elimination of the influence of the rows (= glycine concentration) is obtained:

$$\chi^2_3 = \frac{12}{4 \cdot 8 \cdot 5} (9^2 + 15^2 + 24^2 + 32^2) - 3 \cdot 8(5) = 22.95$$

which confirms the above result as was to be expected in view of the consistency of the table.

Table XXVI

Row ranks of uptake rates

	0 ‰	10 ‰	20 ‰	30 ‰	
20	1 1	2 2	3 3	4 4	
50	1 2	2 1	3 3	4 4	
120	1 1	2 2	3 3	4 4	
240	1 1	2 2	3 3	4 4	
R_i	9	15	24	32	$80 = 8 \frac{(1+4)4}{2}$

2.7.2.5 Examples of multiple *a posteriori* comparisons

In the analysis of variance of Table VIII we got to the conclusion that the 8 samples differ significantly between each other, because

$$F_{7,62} = \frac{.099772}{.011891} = 8.39 > F_{7,62(.95)} = 2.17 \ .$$

It would be interesting to know if all sample means are widely spread or if only some differ particularly from the others.

We first follow the Student-Newman-Keuls procedure, building Table XXVII like (2.3-28).

Table XXVII

Between mean differences

Rank i	Means $\bar{x}_{(i)}$	Rank g	1	2	3	4	5	6	7	8
		Means $\bar{x}_{(g)}$	.069	.072	.080	.123	.204	.209	.221	.241
		No.of individuals n_i n_g	10	5	21	5	5	5	10	10
1	.069	10	-							
2	.072	5	.003	-						
3	.080	21	.011	.008	-					
4	.123	5	.054	.051	.043	-				
5	.204	5	.135	.132	.124	.081	-			
6	.209	5	.140	.137	.129	.086	.005	-		
7	.221	10	.152	.149	.141	.098	.017	.012	-	
8	.241	10	.172	.169	.161	.118	.037	.032	.020	-

From Table VIII we get $MS_w = .011891$ with 62 DF. To apply (2.3-29) we need from Table A-12 the studentized ranges $Q_{k,62(.95)}$ for $k = 2 \ldots 8$, which are:

2.828, 3.398, 3.735, 3.975, 4.161, 4.312, 4.439

which when multiplied by $\sqrt{MS_w} = .109$ become:

.308, .370, .407, .433, .454, .470, .484 .

To test any particular range between two means of Table XXVII, we need only compare it with these quantities multiplied by $\sqrt{(n_i+n_g)/2n_in_g}$. The significant differences of Table XXVII are underlined, thus leading to the conclusion that either the sample means .069, .080 and/or the means .221 and .241 are those who affect mainly the sample MS.

Another approach is that based on the construction of a dendrogram (Sokal and Sneath, 1963), e.g. by single linkage, starting from the differences in Table XXVII.

At the levels of .003 means $\bar{x}_{(1)}$ and $\bar{x}_{(2)}$ join to a cluster, at the level of .005 $\bar{x}_{(5)}$ and $\bar{x}_{(6)}$ join, at the level of .008 $\bar{x}_{(3)}$ joins to $\bar{x}_{(1)}$ and $\bar{x}_{(2)}$ and so on, leading to the dendrogram of Figure 8.

If we evaluate for each cluster the F-ratio with the same MS_w, the only ratio which is higher than the critical value is achieved when clusters with means $[\bar{x}_{(1)}, \bar{x}_{(2)}, \bar{x}_{(3)}, \bar{x}_{(4)}]$ and $[\bar{x}_{(5)}, \bar{x}_{(6)}, \bar{x}_{(7)}, \bar{x}_{(8)}]$ join to a single cluster. Hence, we may conclude, in contrast to the Student-Newman-Keuls procedure that $[\bar{x}_{(1)} \ldots \bar{x}_{(4)}]$ form a homogeneous cluster, so as $[\bar{x}_{(5)} \ldots \bar{x}_{(8)}]$, but that both clusters differ significantly between them. The significance level is obviously not equal to that chosen for the F-ratio since the decision procedure is sequential. It is seen that the dendrogram gives a very good insight of the structure of the between samples variability.

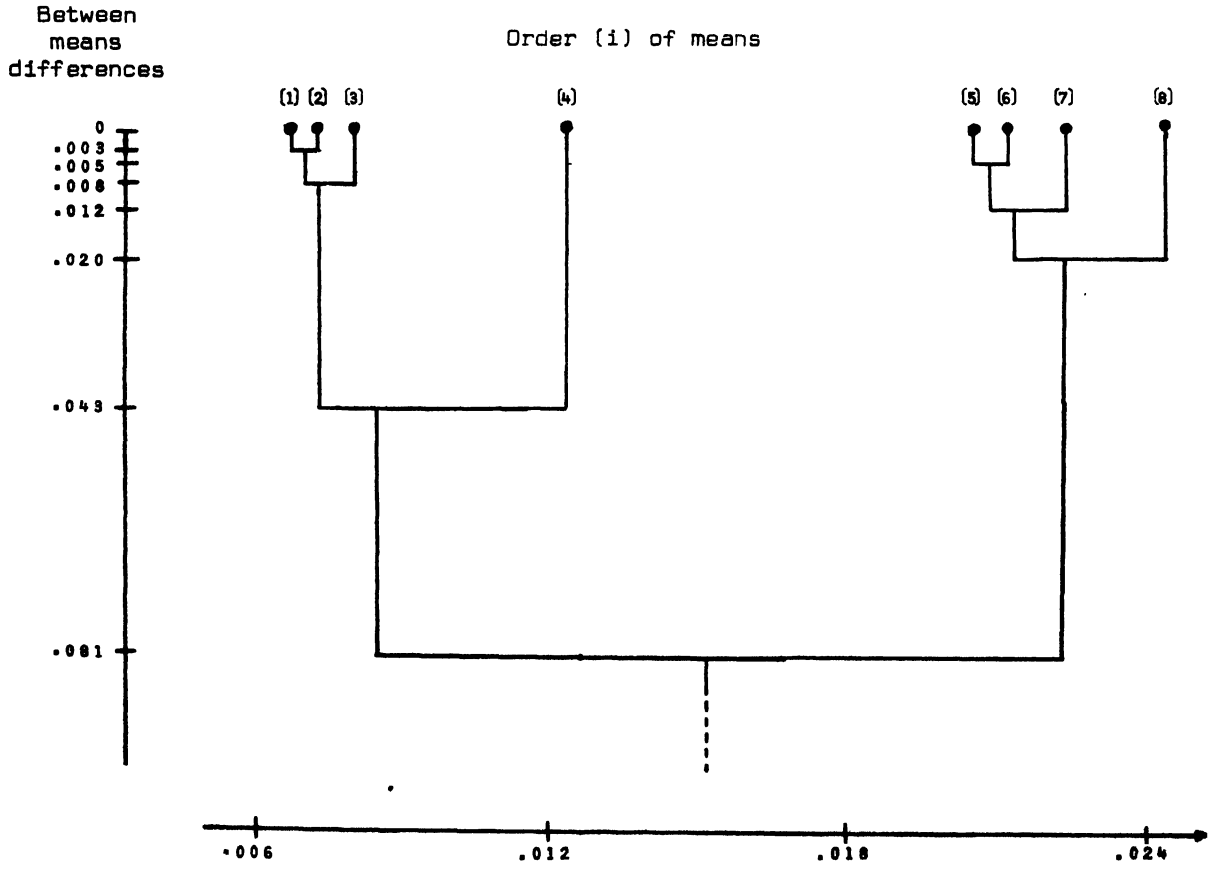


Figure 8 Dendrogram of the means

2.7.3 Examples of tests about variances

The comparison of two variances by the F-ratio (2.4-6) has already been carried out several times as for example in Section 2.7.2.2, and hence will not be repeated here.

An example for the comparison of nine variances by the Cochran test can be found in Section 1.7.1. We will now see an example of Bartlett's and Scheffé's tests.

Let us consider again the nine samples of Table V whose variances are for the first pretreatment and determination:

Sample n^0	1	2	3	4	5	6	7	8	9
Sample size n_i	10	10	5	10	5	5	5	10	21
$10^4 \times$ variances s_i^2	443.67	99.73	9.47	5.16	8.94	39.54	82.25	209.59	19.74

The pooled variance is:

$$10^4 \times s^2 = \frac{\Sigma(n_i-1)s_i^2}{\Sigma n_i - 9} = \frac{7778.95}{72} = 108.04$$

hence,

$$s^2 = .0108 \ .$$

To obtain Bartlett's test (2.4-10) we first evaluate

$$k = 1 + \frac{1}{3 \cdot 8} \left(1.494 - \frac{1}{72}\right) = 1.0617$$

and then

$$\chi_o^2 = 64.24 \ ,$$

which is very high denoting that the samples are heteroscedastic, confirming the impression we got in Section 1.7.4: Probably some outliers or some wrong determinations among the observations have influenced the values of some variances.

Another method of testing the hypothesis of homoscedasticity is that proposed by Scheffé. We have divided randomly the first 8 samples into 2 sub-groups each, and the 9th sample into 3 sub-groups. For each sub-group we have evaluated the variances and their normalizing transformation (2.4-13) obtaining Table XXVIII. Then an ordinary ANOVA has been performed (see Section 1.4.1) obtaining $SS(\theta_i|\mu) = 90.648$, $SS(\varepsilon) = 7.211$ so that $F_{8,10} = 15.71$ which is highly significant confirming the former results.

Table XXVIII

Normalizing transformation of variances × 10^6

	Sample Nos.								
	1	2	3	4	5	6	7	8	9
	8.653 7.007	7.252 7.236	2.292 3.447	3.998 5.340	1.064 3.359	2.944 3.274	3.089 4.332	7.962 6.850	7.817 7.624 7.184
Means	7.8300	7.2440	2.8695	4.6690	2.2115	3.109	3.7105	7.406	7.5417
	$\bar{x}$ = 5.3008								

2.7.4 Examples of tests about regression and correlation

2.7.4.1 Example of a simple regression problem

Fishes placed in a rotating tube compensate for torque until a certain critical velocity is reached. The fish are then forced to rotate with the tube. The critical rate of rotation depends on the age (length) of the fish and of course on its health conditions. Lindahl and Schwanbom (1971) demonstrated the dependence in form of a linear regression of the critical rate of rotation per minute (r.p.m.) on the M.Mercury -OH content (mg/g.w.w.) of *Leuciscus leuciscus*, having obtained the following results:

y_i	r.p.m.	60	60	58	50	52	55	45	58	45	22	25	27	36	39	47	23
x_j	mg/g.w.w.	49	55	70	110	90	85	350	100	260	660	480	460	420	436	300	670

The regression equation $\phi(x_j)$ is given by:

$$\phi(x_j) = 61.333 - .0608\, x_j$$

with a residual variance $S_R^2 = 13.98$. Test t (2.5-23) gives a very high value, according to a determination coefficient of .925, so that the existence of a regression is out of any doubt. To test whether the departures of the observed values y_j from the regression line $\phi(x_j)$ (≡ prevision errors, ε_j) are randomly distributed, we can recur to the run test (2.2-27). Ordering the x_j values we obtain:

x_j	49	55	70	85	90	100	110	260	300	350	420	436	460	480	660	67
ε_j	+1.65	+2.01	+.92	-1.17	-3.86	+2.75	-4.65	-.53	+3.91	+4.95	+.20	+4.18	-6.37	-7.15	+.80	+2.

The sequence of the errors $\varepsilon_j = y_j - \phi(x_j)$ can be considered as a sequence of positive and negative signs, hence as a sequence of 7 runs. Table A-9 gives us for $n_1 = 6$ ($\equiv$ number of negative signs) and $n_2 = 10$ ($\equiv$ number of positive signs) the acceptance interval ($\alpha = .05$) of 4;13, where our computed number of runs is comprised. Hence the departures of the regression line can be assumed to be randomly distributed and the regression line accepted as satisfactory.

Regression and correlation methods are widely used tools and further examples can be found in Sections 1.7.2, 1.7.5 and 1.7.6

2.7.4.2 Example of a rank regression problem

For simplicity, consider a sample of 10 *Scorpaena porcus* of example (1.7.6) for 1973, randomly chosen for j = 34, 37, 40, 7, 4, 11, 42, 3, 22, 30. Applying Grubbs' test for outliers (2.6-5), the values of the last sample (j = 42) must be suspected as outliers. Wanting to know the correlation between the mercury content y_j and the body weight w_j, without loosing the information of this last element, we can apply Spearman's (2.5-9) or Kendall's (2.5-13) rank correlation coefficients. Herefore we write the ranks of one variable (say y) in increasing order and the corresponding ranks of the other variable, w, below them. To obtain Spearman's r_s we evaluate then the sum of the squared differences between the corresponding ranks, whereas for Kendall's τ we count the number of all higher ranks on the right of the ranks of the non-ordered variable (w), obtaining the results summarized as follows:

j	1	2	3	4	5	6	7	8	9	10	Sums
R_{y_j}, Ranks of y	1	2	3	4	5	6	7	8	9	10	
R_{w_j}, Ranks of w	1	2	4	3	5	8	6	7	9	10	
$(R_{y_j} - R_{w_j})^2$	0	0	1	1	0	4	1	1	0	0	8
$S_i = N(R_{w_k} > R_{w_j})_{k>j}$	9	8	6	6	5	2	3	2	1	0	42 =

With the sums in the last columns we can now easily evaluate

$$r_s = 1 - \frac{6}{10 \cdot 99} \cdot 8 = .95$$

and,

$$\tau = \frac{4}{10 \cdot 9} \cdot 42 - 1 = .87 \ .$$

By inspection of Table A-14 we see that both tests are significant at the 1% level, hence the correlation between body weight and mercury content is significant.

2.7.5 Examples of miscellaneous tests

2.7.5.1 Examples of tests for outliers

We first analyse the first determinations of the 9 samples of Table V. Are there any outliers? Applying Dixon's (2.6-2) and Grubbs' test (2.6-5) we get the same response: individual n^0 9 of sample 1 is an outlier since

$$r_{10} = \frac{x_{(10)} - x_{(9)}}{x_{(10)} - x_{(1)}} = \frac{.766 - .266}{.766 - .037} = .686$$

and,

$$T_{10} = \frac{x_{(10)} - \bar{x}}{s} = \frac{.766 - .206}{.211} = 2.65$$

which are both larger than the critical values of Table A-15 and A-16.

For sample 9, being $n_9 > 20$ we could also evaluate (2.6-6):

$$Z_{21} = \frac{R}{S} = \frac{.165}{.044} = 3.75$$

which lies between the critical values of Table A-17, hence confirming that no value must be suspected to be an outlier.

Another example can be obtained by McMahon and Rigler (1962) who examined the heart-beat rate of a *Daphnia* under control conditions, immediately after placing the animals in the observation chamber and after other 10-time intervals, obtaining: 265, 204, 188, 187, 187, 189, 184, 190, 188, 182, 185. Both Dixon and Grubbs' test detect the first 2 values as outliers; hence we may conclude that the heart-beat rate has been normalized after 2 time intervals.

Another example has been seen in Section 2.7.4.2.

2.7.5.2 Examples of tests about randomness

We have already seen an example of the control of randomness in Section 2.7.4.1. Now we will see two applications obtained by Hoppenheit (1977).

A *Tisbe holothuria* culture (Copepoda, Harpacticoida) was permanently exposed to 148 µg Cd/l and exploited weekly (30% removal). The log transformed population densities (per 10 ml) assessed before the weekly harvests, are (Figure 9):

2.4362	2.5145	2.3636	2.5172	2.2900	2.4314	2.1818	2.4857	2.3747
2.4502	2.4564	2.5119	2.0294	2.3802	2.6599	2.5340	2.1673	2.3483
2.3802	2.2122	2.6955	2.3032	2.2718	2.2553	2.0128	2.3892	2.4728

We apply the mean square successive difference test (2.6-7), the circular serial correlation coefficient of lag h = 1 (2.6-8) and the runs up and down test (2.6-9) for which the values of the differences between successive population densities are first transformed into + and - signs:

	+	-	+	-	+	-	+	-
+	+	+	-	+	+	-	-	+
+	-	+	-	-	-	-	+	+

we get the following results:

$$\eta^2_{27} = 2.1889 \quad ; \quad R_{1,27} = -.094 \quad ; \quad z = 3.94$$

It may be underlined that the last test would be exactly the same even without log - transformations.

The results are contradictory. The first 2 parametric tests lie well within the confidence limits of Table A-18 hence leading to the acceptance of the hypothesis of randomness. Instead the third test for runs up and down tells us that there is no trend (because of the positive sign) but the fluctuations are too frequent to be considered randomly distributed. This depends obviously on the non-parametric character of the test which does not make any difference between small and large differences. Since the distribution of the values can be assumed to be normal, and since we are mainly interested to test if there is a trend or not, we may rely on the first two tests. The sign test of Cox and Stuart (Section 2.6.2) confirms no trend.

We consider the same experiment for which weekly 10% were removed, obtaining (Figure 10):

2.1903	2.3075	2.6031	2.6532	2.7388	2.3997	2.4378	2.4742	2.5145
2.2833	2.0453	2.3284	2.5185	2.4564	2.3856	2.3464	2.2967	2.4330
2.1987	1.8062	2.1106	2.5276	1.8865	1.4472	2.0607		

with the signs of the successive differences:

	+	+	+	+	-	+	+	+
-	-	+	+	-	-	-	-	+
-	-	+	+	-	-	+		

$$\eta^2_{25} = .9774 \quad ; \quad R_{1,25} = -.805 \quad ; \quad z = -2.63$$

All tests lie under the lowest critical value and lead to the conclusion:

i) The data are not randomly distributed because of the presence of a trend (η^2, $|R|$, negative sign of z).

ii) The trend is decreasing (sign of R).

Only the sign test of Cox and Stuart with 3 positive signs of 12 does not reach the 5% significance level. But this is mainly due to the fact that this test is based on the differences, hence on half as much values as the other tests, so that the sampling variance naturally increases.

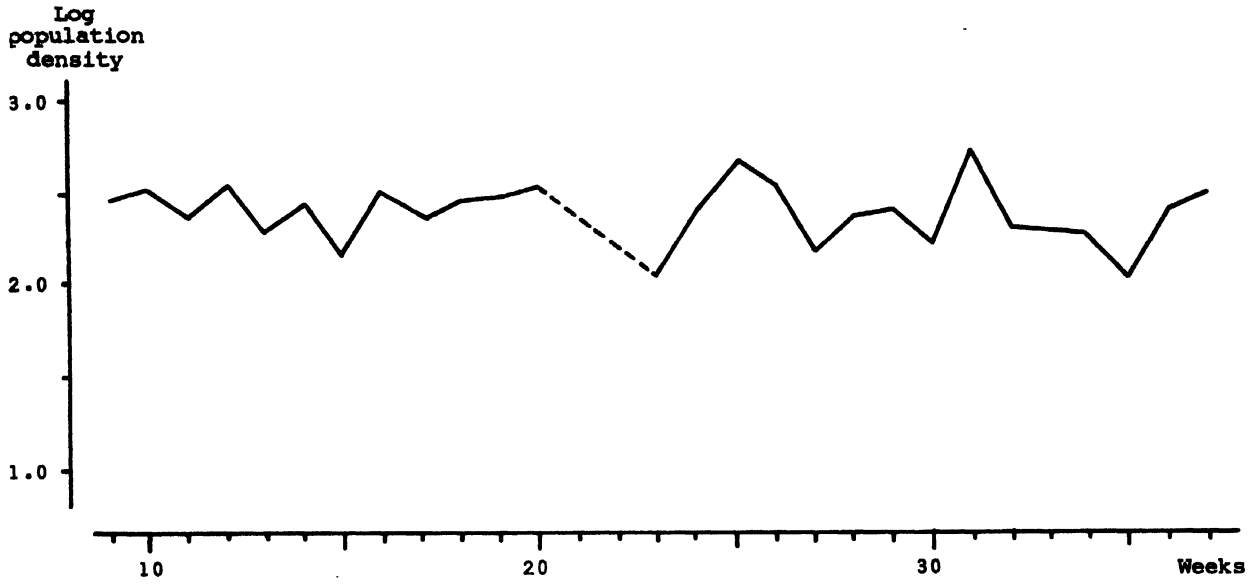


Figure 9 Time sequence of the log transformed population densities of a *Tisbe holothuriae* culture (Copepoda, Harpactoida) with 30% weekly removal

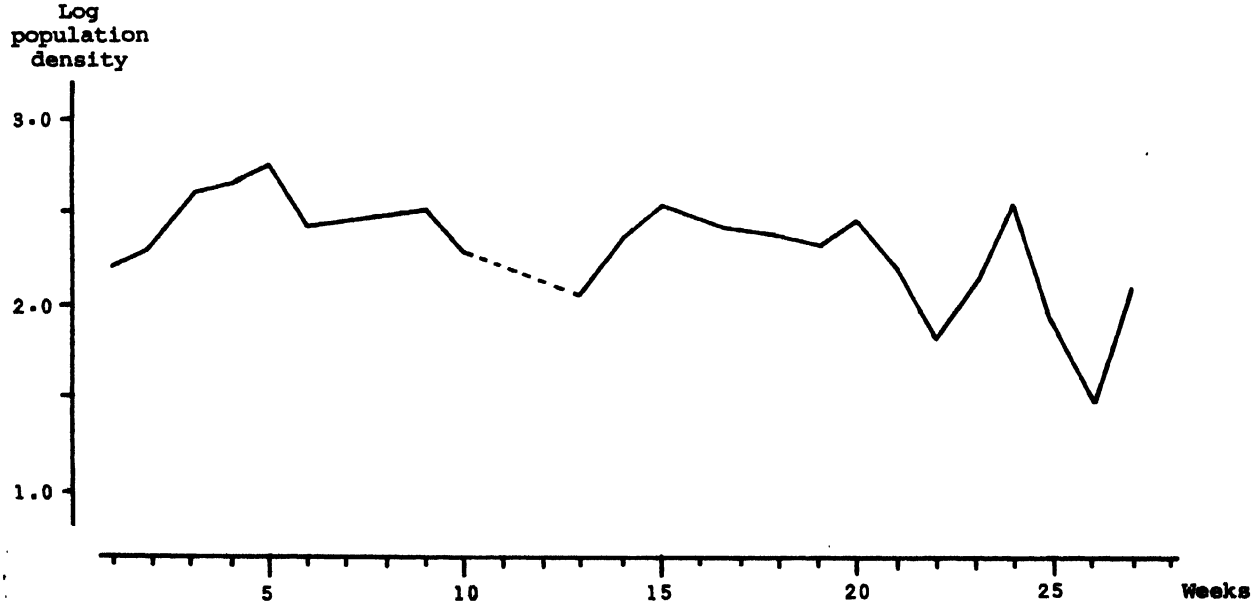


Figure 10 Time sequence of the log transformed population densities of a *Tisbe holothuriae* culture (Copepoda, Harpactoida) with 10% weekly removal

3. REFERENCES

d'Agostino, R. and E.S. Pearson 1973 Tests for departure from normality. Empirical results for the distributions of b_2 and $\sqrt{b_1}$. Biometrika, 60:613-22

Aspin, A.A. 1949 Tables for use in comparisons whose accuracy involves two variances, separately estimated. Biometrika, 36:290-3

Bayne, B.L., R.L. Thomson and J. Wieblows 1973 Some effects of temperature and food on the rate of oxygen consumption by *Mytilus edulis*. In Effects of temperature on ectothermic organisms, edited by E. Wieser. Berlin, Springer, pp.118-93

Bernhard, M. 1976 Manual of methods in aquatic environment research. Part 3. Sampling and analysis of biological material (Guidelines for the FAO(GFCM)/UNEP Joint Co-ordinated Project on Pollution in the Mediterranean). FAO Fish.Tech.Pap., (158):124 p.

Bowman, K.O. and L.R. Shenton 1975 Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2. Biometrika, 62:243-9

Bulnheim, H.-P. 1977 Effects of inbreeding on the relative fitness of the amphipod *Gammarus duebeni*, Liljeborg 1852. Crustaceana (Suppl.4):3-14

Cochran, W.G. 1941 The distribution of the largest of a set of estimated variances as a fraction of their total. Ann.Eugen., 11:47-52

_____ 1960 Sampling techniques. New York, Wiley, 330 p.

Conover, W.J. 1971 Practical non-parametric statistics. New York, Wiley, 462 p.

Cox, D.R. and A. Stuart 1955 Some quick tests for trend in location and dispersion. Biometrika, 42:80-95

David, F.N. 1954 Tables of the ordinates and probability integral of the distribution of the correlation coefficient in small samples. London, Cambridge University Press, 57 p.

David, H.A. 1952 Upper 5 and 1% points of the maximum F-ratio. Biometrika, 39:422-4

David, H.A., H.O. Hartley and E.S. Pearson 1954 The distribution of the ratio, in a single normal sample, of range to standard deviation. Biometrika, 41:482-93

Diem, K. (ed.) 1962 Documenta Geigy; Scientific Tables. Basel, J.R. Geigy, 742 p.

Dixon, W.J. 1951 Ratios involving extreme values. Ann.Math.Statist., 22:68-78

Dunn, O.J. and V.A. Clark 1974 Applied statistics. New York, Wiley, 387 p.

Elliott, J.M. 1971 Statistical analysis of samples of benthic invertebrates. Sci.Publ. Freshwat.Biol.Assoc., (25):144 p.

Filliben, J.J. 1975 The probability plot correlation coefficient test for normality. Technometrics, 17:111-7

Finney, D.J. 1971 Statistical method in biological assay. High Wycombe and London, C. Griffin, 668 p.

Fisher, R.A. 1954 Statistical methods for research workers. Edinburgh, Oliver & Boyd, 12th ed:356 p.

Fisher, R.A. and F. Yates 1963 Statistical tables for biological, agricultural and medical research. Edinburgh, Oliver & Boyd, 6th ed:138 p.

Friedman, M. 1937 The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J.Am.Statist.Assoc., 32:675-701

_____ 1940 A comparison of alternative tests of significance for the problem of m rankings. Ann.Math.Statist., 11:86-92

Gabriel, K.R. 1964 A procedure for testing the homogeneity of all sets of means in analysis of variance. Biometrics, 20:459-77

Garner, W.R. and W.J. McGill 1956 The relation between information and variance analysis. Psychometrika, 21:219-28

Grubbs, F.E. 1969 Procedures for detecting outlying observations in samples. Technometrics, 11:1-21

Harter, H.L. 1960 Tables of range and studentized range. Ann.Math.Statist., 31:1122-47

Hatfield, C.T. and P.H. Johansen 1972 Effects of four insecticides on the ability of Atlantic salmon parr *(Salmo salar)* to learn and retain a simple conditioned response. J.Fish.Res.Board Can., 29:315-21

Hoppenheit, M. 1977 On the dynamics of exploited populations of *Tisbe holothuriae* (Copepoda, Harpacticoida). Helgol.Wiss.Meeresunters., 29:503-23

Kendall, M.G. and A. Stuart 1961 The advanced theory of statistics. London, Ch. Griffin, vol.2:676 p.

Kolmogoroff, A. 1941 Confidence limits for an unknown distribution function. Ann.Math. Statist., 12:461-3

Kruskal, W.H. and W.A. Wallis 1952 Use of ranks in one-criterion variance analysis. J.Am.Statist.Assoc., 47:583-621

1953 Use of ranks in one-criterion variance analysis. Addendum. J.Am.Statist.Assoc., 48:907-11

Kullback, S. 1959 Information theory and statistics. New York, Wiley, 395 p.

Labat, R., J. Pequignot and A. Chatelet 1974 Action toxique du cuivre sur les branchies de carpe. Ann.Limnol., 10:109-14

Larsson, A. 1975 Some biochemical effects of cadmium on fish. In Sublethal effects of toxic chemicals on aquatic animals, edited by J.H. Koeman and J.J.T.W.A. Strik. Amsterdam, Elsevier, pp. 3-13

Lindahl, P.E. and E. Schwanbom 1971 Rotary flow techniques as a means of detecting sublethal poisoning in fish populations. Oikos, 22:354-7

Lord, E. 1956 The use of range in place of standard deviation in the t-test. Biometrika, 34:41-67

Maitland, P.S. and M.M. Penney 1967 The ecology of the Simuliidae in a Scottish river. J.Anim.Ecol., 36:179-206

McMahon, J.W. and F.H. Rigler 1962 Mechanisms regulating the feeding rate of *Daphnia magna* Straus. Can.J.Zool., 41:321-32

McNemar, Q. 1962 Psychological statistics. New York, Wiley

Merrington, M. and C.M. Thompson 1943 Tables of percentage points of the inverted beta (F) distribution. Biometrika, 33:73-88

Miller, L.H. 1956 Table of percentage points of Kolmogoroff statistics. J.Am.Statist.Assoc., 51:111-21

Möller, F. 1961 Tabelle tetracoriche - una precisazione sul método esatto di R.A. Fisher. Boll.Ric.Oper., (5):33-5

1973 Problemi di inferenza nel modello lineare. Milano, Cisalpina-Goliardica, 162 p.

1976 Basic statistical methods in biological assays. In Lectures presented at the fourth FAO/SIDA Training Course on Aquatic Pollution in Relation to Protection of Living Resources. Rome, FAO,TF-INT/73(SWE)-Suppl.1:211-31

Neumann, J. 1941 Distribution of the ratio of the mean square successive difference to the variance. Ann.Math.Statist., 12:367-95

Owen, D.B. 1962 Handbook of statistical tables. Reading, Mass., Addison-Wesley, 349 p.

Pearson, E.S. 1930 A further development of tests for normality. Biometrika, 22:239-49

Pearson, E.S. and H.O. Hartley 1958 Biometrika tables for statisticians. 1. London, Cambridge University Press, 2nd ed:240 p.

Rao, C.R. 1952 Advanced statistical methods in biometric research. New York, Wiley, 390 p.

Renzoni, A. 1977 A case of mercury abatement along the Tuscan coast. In IIIes journées d'études sur les pollutions marines, Split 1976. Monaco, CIESMM. pp.95-7

Rohlf, F.J. and R.R. Sokal 1969 Statistical tables. San Francisco, W.H. Freeman and Co., 253 p.

Sachs, L. 1970 Statistische Methoden, ein Soforthelfer. Berlin, Springer, 103 p. and tables

Scheffé, H. 1959 The analysis of variance. New York, Wiley, 477 p.

Searle, S.R. 1971 Linear models. New York, Wiley, 532 p.

Shapiro, S.S. and M.B. Wilk 1965 An analysis of variance test for normality (complete samples). Biometrika, 52:591-611

Shapiro, S.S., M.B. Wilk and H.J. Chen 1968 A comparative study of various tests of normality. J.Am.Statist.Assoc., 63:1343-72

Shenton, L.R. and K.O. Bowman 1977 A bivariate model for the distribution of $\sqrt{b_1}$ and b_2. J.Am.Statist.Assoc., 72:206-11

Siebers, D. and H.-P. Bulnheim 1977 Salinity dependence, uptake kinetics, and specificity of amino-acid absorption across the body surface of the oligochaete annelid *Enchytraeus albidus*. Helgol.Wiss.Meeresunters., 29:473-92

Smirnov, N. 1948 Table for estimating the goodness of fit of empirical distributions. Ann.Math.Statist., 19:279-81

Snedecor, G.W. 1956 Statistical methods. Ames, Iowa State College Press, 5th ed:534 p.

Sokal, R.R. and F.J. Rohlf 1969 Biometry. San Francisco, W.H. Freeman and Co., 776 p.

Sokal, R.R. and P.H.A. Sneath 1963 Principles of numerical taxonomy. San Francisco, Freeman Co., 359 p.

Stange, K. and H.J. Henning 1966 Graf/Henning/Stange - Formeln und Tabellen der mathematischen Statistik. Berlin, Springer, 362 p.

Stöppler, M. et al. Mercury in marine organisms in the Mediterranean and other European seas. Rapp.P.-Réun.CIESMM, 24 (in press)

Store, G. 1972 Contribution a l'étude de la notion de concentration lethale limite moyenne (CL 50) appliquée a des invertébrés marins. 1. Etude méthodologique. Téthys, 4:597-644

Swed, F.S. and C. Eisenhart 1943 Tables for testing randomness of grouping in a sequence of alternatives. Ann.Math.Statist., 14:66-87

Thompson, C.M. 1941 Tables of percentage points of the χ^2 distribution. Biometrika, 32: 187-91

Van der Waerden, B.L. 1971 Mathematische Statistik. Berlin, Springer, 360 p.

Wald, A. and J. Wolfowitz 1940 On a test whether two samples are from the same population. Ann.Math.Statist., 11:147-62

Wilcoxon, F. and R.A. Wilcox 1964 Some rapid approximate statistical procedures. New York, American Cyanamid Co., 28 p.

ANNEX

STATISTICAL TABLES

Table A-1

Common logarithms of factorials
(from Rohlf and Sokal, 1969, with kind permission of W.H. Freeman and Co., San Francisco ©)

	0	10	20	30	40
0	0.0000	6.5598	18.3861	32.4237	47.9116
1	0.0000	7.6012	19.7083	33.9150	49.5244
2	0.3010	8.6803	21.0508	35.4202	51.1477
3	0.7782	9.7943	22.4125	36.9387	52.7811
4	1.3802	10.9404	23.7927	38.4702	54.4246
5	2.0792	12.1165	25.1906	40.0142	56.0778
6	2.8573	13.3206	26.6056	41.5705	57.7406
7	3.7024	14.5511	28.0370	43.1387	59.4127
8	4.6055	15.8063	29.4841	44.7185	61.0939
9	5.5598	17.0851	30.9465	46.3096	62.7841

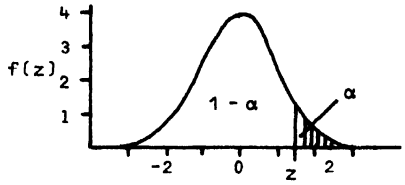
Table A-2

Correction factors for ranges c_n, d_n and g_n and critical values for the median h_n
(after Stange and Henning, 1966, and Sachs, 1970,
with kind permission of Springer Verlag, Berlin)

n	c_n	d_n	g_n	h_n
2	1.000	1.128	.756	-
3	1.160	1.693	.525	-
4	1.092	2.059	.427	-
5	1.198	2.326	.371	-
6	1.136	2.534	.335	0
7	1.214	2.704	.308	0
8	1.159	2.847	.288	0
9	1.223	2.970	.272	1
10	1.175	3.078	.259	1
11	1.229	3.173	.248	1
12	1.190	3.258	.239	2
13	1.233	3.336	.231	2
14	1.195	3.407	.224	2
15	1.237	3.472	.217	3
16	1.202	3.532	.212	3
17	1.238	3.588	.207	4
18	1.207	3.640	.203	4
19	1.239	3.689	.199	4
20	1.212	3.735	.195	5

Table A-3

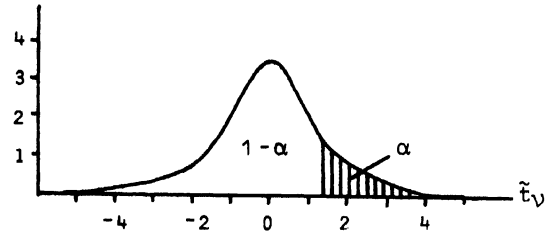
Critical values of the normal variate z
$(-z_{(\alpha)} = z_{(1-\alpha)})$
(from Diem, 1962, with kind permission of J.R. Geigy, Basel)



α→ ↓	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.50	0.0000	0.0025	0.0050	0.0075	0.0100	0.0125	0.0150	0.0175	0.0201	0.0226
0.51	0.0251	0.0276	0.0301	0.0326	0.0351	0.0376	0.0401	0.0426	0.0451	0.0476
0.52	0.0502	0.0527	0.0552	0.0577	0.0602	0.0627	0.0652	0.0677	0.0702	0.0728
0.53	0.0753	0.0778	0.0803	0.0828	0.0853	0.0878	0.0904	0.0929	0.0954	0.0979
0.54	0.1004	0.1030	0.1055	0.1080	0.1105	0.1130	0.1156	0.1181	0.1206	0.1231
0.55	0.1257	0.1282	0.1307	0.1332	0.1358	0.1383	0.1408	0.1434	0.1459	0.1484
0.56	0.1510	0.1535	0.1560	0.1586	0.1611	0.1637	0.1662	0.1687	0.1713	0.1738
0.57	0.1764	0.1789	0.1815	0.1840	0.1866	0.1891	0.1917	0.1942	0.1968	0.1993
0.58	0.2019	0.2045	0.2070	0.2096	0.2121	0.2147	0.2173	0.2198	0.2224	0.2250
0.59	0.2275	0.2301	0.2327	0.2353	0.2378	0.2404	0.2430	0.2456	0.2482	0.2508
0.60	0.2533	0.2559	0.2585	0.2611	0.2637	0.2663	0.2689	0.2715	0.2741	0.2767
0.61	0.2793	0.2819	0.2845	0.2871	0.2898	0.2924	0.2950	0.2976	0.3002	0.3029
0.62	0.3055	0.3081	0.3107	0 3134	0.3160	0.3186	0.3213	0.3239	0.3266	0.3292
0.63	0.3319	0.3345	0.3372	0.3398	0.3425	0.3451	0.3478	0.3505	0.3531	0.3558
0.64	0.3585	0.3611	0.3638	0.3665	0.3692	0.3719	0.3745	0.3772	0.3799	0.3826
0.65	0.3853	0.3880	0.3907	0.3934	0.3961	0.3989	0.4016	0.4043	0.4070	0.4097
0.66	0.4125	0.4152	0.4179	0.4207	0.4234	0.4261	0.4289	0.4316	0.4344	0.4372
0.67	0.4399	0.4427	0.4454	0.4482	0.4510	0.4538	0.4565	0.4593	0.4621	0.4649
0.68	0.4677	0.4705	0.4733	0.4761	0.4789	0.4817	0.4845	0.4874	0.4902	0.4930
0.69	0.4959	0.4987	0.5015	0.5044	0.5072	0.5101	0.5129	0.5158	0.5187	0.5215
0.70	0.5244	0.5273	0.5302	0.5330	0.5359	0.5388	0.5417	0.5446	0.5476	0.5505
0.71	0.5534	0.5563	0.5592	0.5622	0.5651	0.5681	0.5710	0.5740	0.5769	0.5799
0.72	0.5828	0.5858	0.5888	0.5918	0.5948	0.5978	0.6008	0.6038	0.6068	0.6098
0.73	0.6128	0.6158	0.6189	0.6219	0.6250	0.6280	0.6311	0.6341	0.6372	0.6403
0.74	0.6433	0.6464	0.6495	0.6526	0.6557	0.6588	0.6620	0.6651	0.6682	0.6713
0.75	0.6745	0.6776	0.6808	0.6840	0.6871	0.6903	0.6935	0.6967	0.6999	0.7031
0.76	0.7063	0.7095	0.7128	0.7160	0.7192	0.7225	0.7257	0.7290	0.7323	0.7356
0.77	0.7388	0.7421	0.7454	0.7488	0.7521	0.7554	0.7588	0.7621	0.7655	0.7688
0.78	0.7722	0.7756	0.7790	0.7824	0.7858	0.7892	0.7926	0.7961	0.7995	0.8030
0.79	0.8064	0.8099	0.8134	0.8169	0.8204	0.8239	0.8274	0.8310	0.8345	0.8381
0.80	0.8416	0.8452	0.8488	0.8524	0.8560	0.8596	0.8633	0.8669	0.8705	0.8742
0.81	0.8779	0.8816	0.8853	0.8890	0.8927	0.8965	0.9002	0.9040	0.9078	0.9116
0.82	0.9154	0.9192	0.9230	0.9269	0.9307	0.9346	0.9385	0.9424	0.9463	0.9502
0.83	0.9542	0.9581	0.9621	0.9661	0.9701	0.9741	0.9782	0.9822	0.9863	0.9904
0.84	0.9945	0.9986	1.0027	1.0069	1.0110	1.0152	1.0194	1.0237	1.0279	1.0322
0.85	1.0364	1.0407	1.0450	1.0494	1.0537	1.0581	1.0625	1.0669	1.0714	1.0758
0.86	1.0803	1.0848	1.0893	1.0939	1.0985	1.1031	1.1077	1.1123	1.1170	1.1217
0.87	1.1264	1.1311	1.1359	1.1407	1.1455	1.1503	1.1552	1.1601	1.1650	1.1700
0.88	1.1750	1.1800	1.1850	1.1901	1.1952	1.2004	1.2055	1.2107	1.2160	1.2212
0.89	1.2265	1.2319	1.2372	1.2426	1.2481	1.2536	1.2591	1.2646	1.2702	1.2759
0.90	1.2816	1.2873	1.2930	1.2988	1.3047	1.3106	1.3165	1.3225	1.3285	1.3346
0.91	1.3408	1.3469	1.3532	1.3595	1.3658	1.3722	1.3787	1.3852	1.3917	1.3984
0.92	1.4051	1.4118	1.4187	1.4255	1.4325	1.4395	1.4466	1.4538	1.4611	1.4684
0.93	1.4758	1.4833	1.4909	1.4985	1.5063	1.5141	1.5220	1.5301	1.5382	1.5464
0.94	1.5548	1.5632	1.5718	1.5805	1.5893	1.5982	1.6072	1.6164	1.6258	1.6352
0.95	1.6449	1.6546	1.6646	1.6747	1.6849	1.6954	1.7060	1.7169	1.7279	1.7392
0.96	1.7507	1.7624	1.7744	1.7866	1.7991	1.8119	1.8250	1.8384	1.8522	1.8663
0.97	1.8808	1.8957	1.9110	1.9268	1.9431	1.9600	1.9774	1.9954	2.0141	2.0335
0.98	2.0537	2.0749	2.0969	2.1201	2.1444	2.1701	2.1973	2.2262	2.2571	2.2904
0.99	2.3263	2.3656	2.4089	2.4573	2.5121	2.5758	2.6521	2.7478	2.8782	3.0902

Table A-4

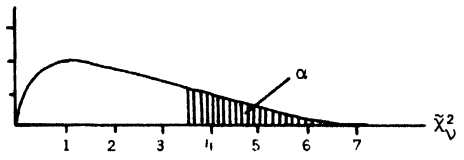
Critical values of the t_ν variate ($-t_{(\alpha)} = t_{(1-\alpha)}$)
(after Diem, 1962, with kind permission of J.R. Geigy, Basel)



ν	α=.05	.025	.01	.005
1	6.314	12.706	31.821	63.657
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.813	2.228	2.764	3.169
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055
13	1.771	2.160	2.650	3.012
14	1.761	2.145	2.624	2.977
15	1.753	2.131	2.602	2.947
16	1.746	2.120	2.583	2.921
17	1.740	2.110	2.567	2.898
18	1.734	2.101	2.552	2.878
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
21	1.721	2.080	2.518	2.831
22	1.717	2.074	2.508	2.819
23	1.714	2.069	2.500	2.807
24	1.711	2.064	2.492	2.797
25	1.708	2.060	2.485	2.787
26	1.706	2.056	2.479	2.779
27	1.703	2.052	2.473	2.771
28	1.701	2.048	2.467	2.763
29	1.699	2.045	2.462	2.756
30	1.697	2.042	2.457	2.750
40	1.684	2.021	2.423	2.704
50	1.676	2.009	2.403	2.678
60	1.671	2.000	2.390	2.660
80	1.664	1.990	2.374	2.639
100	1.660	1.984	2.364	2.626
∞	1.645	1.960	2.326	2.576

Table A-5

Critical values of the χ^2_ν variate
(from Thompson, 1941 (ν = 1-30), and Rohlf and Sokal, 1969, (ν = 31-50), with kind permission of Biometrika Trustees, London, and of W.H. Freeman and Co., San Francisco ©)



ν \ α	0.995	0.975	0.90	0.5	0.1	0.05	0.025	0.01	0.005	ν
1	.000	.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	10
11	2.603	3.816	5.578	10.341	17.275	19.675	21.920	24.725	26.757	11
12	3.074	4.404	6.304	11.240	18.549	21.026	23.337	26.217	28.300	12
13	3.565	5.009	7.042	12.340	19.812	22.362	24.736	27.688	29.819	13
14	4.075	5.629	7.790	13.339	21.064	23.685	26.119	29.141	31.319	14
15	4.601	6.262	8.547	14.339	22.307	24.996	27.488	30.578	32.801	15
16	5.142	6.908	9.312	15.338	23.542	26.296	28.845	32.000	34.267	16
17	5.697	7.564	10.085	16.338	24.769	27.587	30.191	33.409	35.718	17
18	6.265	8.231	10.865	17.338	25.989	28.869	31.526	34.805	37.156	18
19	6.844	8.907	11.651	18.338	27.204	30.144	32.852	36.191	38.582	19
20	7.434	9.591	12.443	19.337	28.412	31.410	34.170	37.566	39.997	20
21	8.034	10.283	13.240	20.337	29.615	32.670	35.479	38.932	41.401	21
22	8.643	10.982	14.042	21.337	30.813	33.924	36.781	40.289	42.796	22
23	9.260	11.688	14.848	22.337	32.007	35.172	38.076	41.638	44.181	23
24	9.886	12.401	15.659	23.337	33.196	36.415	39.364	42.980	45.558	24
25	10.520	13.120	16.473	24.337	34.382	37.652	40.646	44.314	46.928	25
26	11.160	13.844	17.292	25.336	35.563	38.885	41.923	45.642	48.290	26
27	11.808	14.573	18.114	26.336	36.741	40.113	43.194	46.963	49.645	27
28	12.461	15.308	18.939	27.336	37.916	41.337	44.461	48.278	50.993	28
29	13.121	16.047	19.768	28.336	39.088	42.557	45.722	49.588	52.336	29
30	13.787	16.791	20.599	29.336	40.256	43.773	46.979	50.892	53.672	30
31	14.458	17.539	21.434	30.336	41.422	44.985	48.232	52.192	55.003	31
32	15.135	18.291	22.271	31.336	42.585	46.194	49.481	53.486	56.329	32
33	15.816	10.047	23.110	32.336	43.745	47.400	50.725	54.776	57.649	33
34	16.502	19.806	23.952	33.336	44.903	48.602	51.966	56.061	58.964	34
35	17.192	20.570	24.797	34.336	46.059	49.802	53.203	57.342	60.275	35
36	17.887	21.336	25.643	35.336	47.212	50.998	54.437	58.619	61.582	36
37	18.586	22.106	26.492	36.335	48.363	52.192	55.668	59.893	62.884	37
38	19.289	22.879	27.343	37.335	49.513	53.384	56.896	61.162	64.182	38
39	19.996	23.654	28.196	38.335	50.660	54.572	58.120	62.428	65.476	39
40	20.707	24.433	29.051	39.335	51.805	55.758	59.342	63.691	66.766	40
41	21.421	25.215	29.907	40.335	52.949	56.942	60.561	64.950	68.053	41
42	22.139	25.999	30.765	41.335	54.090	58.124	61.777	66.206	69.336	42
43	22.860	26.786	31.625	42.335	55.230	59.304	62.990	67.460	70.616	43
44	23.584	27.575	32.487	43.335	56.369	60.481	64.202	68.710	71.893	44
45	24.311	28.366	33.350	44.335	57.505	61.656	65.410	69.957	73.166	45
46	25.042	29.160	34.215	45.335	58.641	62.830	66.617	71.202	74.437	46
47	25.775	29.956	35.081	46.335	59.774	64.001	67.821	72.443	75.704	47
48	26.511	30.755	35.949	47.335	60.907	65.171	69.023	73.683	76.969	48
49	27.250	31.555	36.818	48.335	62.038	66.339	70.222	74.920	78.231	49
50	27.991	32.357	37.689	49.335	63.167	67.505	71.420	76.154	79.490	50

For values of $\nu > 50$ the $\chi^2_{\nu(1-\alpha)}$ may be normalized by the transformation

$$z = \left[\left(\frac{\chi^2_\nu}{\nu}\right)^{1/3} + \frac{2}{9n} - 1\right]\sqrt{\frac{9n}{2}}$$

which can be compared with the critical values of $\tilde{z}$.

Table A-6.1

Critical values of the F_{ν_1,ν_2} variate at the .95 confidence level
(from Merrington and Thompson, 1943, with kind permission of Biometrika Trustees, London)

Degrees of freedom in denominator, ν_2	Degrees of freedom in numerator, ν_1									
	1	2	3	4	5	6	8	12	24	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	243.9	249.1	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.85	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.79	2.61	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.85	2.69	2.51	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.02	2.39	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.91	2.70	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.37	2.20	2.01	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.18	1.98	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.34	2.16	1.96	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.31	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.29	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.04	1.79	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.10	1.92	1.70	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83	1.61	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	1.94	1.75	1.52	1.00

Table A-6.2

Critical values of the F_{ν_1,ν_2} variate at the .975 confidence level
(from Merrington and Thompson, 1943, with kind permission of Biometrika Trustees, London)

Degrees of freedom in denominator, ν_2	Degrees of freedom in numerator, ν_1									
	1	2	3	4	5	6	8	12	24	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	956.7	976.7	997.2	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.37	39.41	39.46	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.54	14.34	14.12	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	8.98	8.75	8.51	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.76	6.52	6.28	6.02
6	8.81	7.26	6.60	6.63	5.99	5.82	5.60	5.37	5.12	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.89	4.67	4.42	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.43	4.20	3.95	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.10	3.87	3.61	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.85	3.62	3.37	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.66	3.43	3.17	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.51	3.28	3.02	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.39	3.15	2.89	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.29	3.05	2.79	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.20	2.96	2.70	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.12	2.89	2.63	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.06	2.82	2.56	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.01	2.77	2.50	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	2.76	2.72	2.45	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	2.91	2.68	2.41	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.87	2.64	2.37	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.84	2.60	2.33	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.81	2.57	2.30	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.78	2.54	2.27	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.75	2.51	2.24	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.73	2.49	2.22	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.71	2.47	2.19	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.69	2.45	2.17	1.83
29	5.59	4.20	3.60	3.27	3.04	2.88	2.67	2.43	2.15	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.65	2.41	2.14	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.53	2.29	2.01	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.41	2.17	1.88	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.30	2.05	1.76	1.31
∞	5.02	3.69	3.11	2.79	2.57	2.41	2.19	1.94	1.64	1.00

Table A-6.3

Critical values of the F_{ν_1,ν_2} variate at the .99 confidence level
(from Merrington and Thompson, 1943, with kind permission of Biometrika Trustees, London)

Degrees of freedom in denominator, ν_2	Degrees of freedom in numerator, ν_1									
	1	2	3	4	5	6	8	12	24	∞
1	4052	4999.5	5403	5625	5764	5859	5981	6106	6235	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.37	99.42	99.46	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.49	27.05	26.60	26.13
4	21.20	18.00	16.69	16.98	15.52	15.21	14.80	14.37	13.93	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.29	9.89	9.47	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.31	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.07	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.28	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.73	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.33	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.74	4.40	4.02	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.78	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.30	3.96	3.59	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.14	3.80	3.43	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.67	3.29	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.18	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.79	3.46	3.08	2.65
18	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.00	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.63	3.30	2.92	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.23	2.86	2.42
21	8.02	5.75	4.87	4.37	4.04	3.81	3.51	3.17	2.80	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.45	3.12	2.75	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.41	3.07	2.70	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.36	3.03	2.66	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.32	2.99	2.62	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.29	2.96	2.58	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.26	2.93	2.55	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.23	2.90	2.52	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.20	2.87	2.49	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.47	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.29	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.12	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.66	2.34	1.95	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.51	2.18	1.79	1.00

Table A-7

Critical values of the Kolmogoroff-Smirnov test
(from Miller, 1956, with kind permission of the American Statistical Association, Washington, D.C.)

n \ α	0.2	0.1	0.05	0.02	0.01
1	.90000	.95000	.97500	.99000	.99500
2	.68377	.77639	.84189	.90000	.92929
3	.56481	.63604	.70760	.78456	.82900
4	.49265	.56522	.62394	.68887	.73424
5	.44698	.50945	.56328	.62718	.66853
6	.41037	.46799	.51926	.57741	.61661
7	.38148	.43607	.48342	.53844	.57581
8	.35831	.40962	.45427	.50654	.54179
9	.33910	.38746	.43001	.47960	.51332
10	.32260	.36866	.40925	.45662	.48893
11	.30829	.35242	.39122	.43670	.46770
12	.29577	.33815	.37543	.41918	.44905
13	.28470	.32549	.36143	.40362	.43247
14	.27481	.31417	.34890	.38970	.41762
15	.26588	.30397	.33760	.37713	.40420
16	.25778	.29472	.32733	.36571	.39201
17	.25039	.28627	.31796	.35528	.38086
18	.24360	.27851	.30936	.34569	.37062
19	.23735	.27136	.30143	.33685	.36117
20	.23156	.26473	.29408	.32866	.35241
21	.22617	.25858	.28724	.32104	.34427
22	.22115	.25283	.28087	.31394	.33666
23	.21645	.24746	.27490	.30728	.32954
24	.21205	.24242	.26931	.30104	.32286
25	.20790	.23768	.26404	.29516	.31657
26	.20399	.23320	.25907	.28962	.31064
27	.20030	.22898	.25438	.28438	.30502
28	.19680	.22497	.24993	.27942	.29971
29	.19348	.22117	.24571	.27471	.29466
30	.19032	.21756	.24170	.27023	.28987
31	.18732	.21412	.23788	.26596	.28530
32	.18445	.21085	.23424	.26189	.28094
33	.18171	.20771	.23076	.25801	.27677
34	.17909	.20472	.22743	.25429	.27279
35	.17659	.20185	.22425	.25073	.26897
36	.17418	.19910	.22119	.24732	.26532
37	.17188	.19646	.21826	.24404	.26180
38	.16966	.19392	.21544	.24089	.25843
39	.16753	.19148	.21273	.23786	.25518
40	.16547	.18913	.21012	.23494	.25205
41	.16349	.18687	.20760	.23213	.24904
42	.16158	.18468	.20517	.22941	.24613
43	.15974	.18257	.20283	.22679	.24332
44	.15796	.18053	.20056	.22426	.24060
45	.15623	.17856	.19837	.22181	.23798
46	.15457	.17665	.19625	.21944	.23544
47	.15295	.17481	.19420	.21715	.23298
48	.15139	.17302	.19221	.21493	.23059
49	.14987	.17128	.19028	.21277	.22828
50	.14840	.16959	.18841	.21068	.22604
51	.14697	.16796	.18659	.20864	.22386
52	.14558	.16637	.18482	.20667	.22174
53	.14423	.16483	.18311	.20475	.21968
54	.14292	.16332	.18144	.20289	.21768
55	.14164	.16186	.17981	.20107	.21574
56	.14040	.16044	.17823	.19930	.21384
57	.13919	.15906	.17669	.19758	.21199
58	.13801	.15771	.17519	.19590	.21019
59	.13686	.15639	.17373	.19427	.20844
60	.13573	.15511	.17231	.19267	.20673
61	.13464	.15385	.17091	.19112	.20506
62	.13357	.15263	.16956	.18960	.20343
63	.13253	.15144	.16823	.18812	.20184
64	.13151	.15027	.16693	.18667	.20029
65	.13052	.14913	.16567	.18525	.19877
66	.12954	.14802	.16443	.18387	.19729
67	.12859	.14693	.16322	.18252	.19584
68	.12766	.14587	.16204	.18119	.19442
69	.12675	.14483	.16088	.17990	.19303
70	.12586	.14381	.15975	.17863	.19167
71	.12499	.14281	.15864	.17739	.19034
72	.12413	.14183	.15755	.17618	.18903
73	.12329	.14087	.15649	.17498	.18776
74	.12247	.13993	.15544	.17382	.18650
75	.12167	.13901	.15442	.17268	.18528
76	.12088	.13811	.15342	.17155	.18408
77	.12011	.13723	.15244	.17045	.18290
78	.11935	.13636	.15147	.16938	.18174
79	.11860	.13551	.15052	.16832	.18060
80	.11787	.13467	.14960	.16728	.17949
81	.11716	.13385	.14868	.16626	.17840
82	.11645	.13305	.14779	.16526	.17732
83	.11576	.13226	.14691	.16428	.17627
84	.11508	.13148	.14605	.16331	.17523
85	.11442	.13072	.14520	.16236	.17421
86	.11376	.12997	.14437	.16143	.17321
87	.11311	.12923	.14355	.16051	.17223
88	.11248	.12850	.14274	.15961	.17126
89	.11186	.12779	.14195	.15873	.17031
90	.11125	.12709	.14117	.15786	.16938
91	.11064	.12640	.14040	.15700	.16846
92	.11005	.12572	.13965	.15616	.16755
93	.10947	.12506	.13891	.15533	.16666
94	.10889	.12440	.13818	.15451	.16579
95	.10833	.12375	.13746	.15371	.16493
96	.10777	.12312	.13675	.15291	.16408
97	.10722	.12249	.13606	.15214	.16324
98	.10668	.12187	.13537	.15137	.16242
99	.10615	.12126	.13469	.15061	.16161
100	.10563	.12067	.13403	.14987	.16081

Table A-8

Critical values of the Mann-Whitney-Wilcoxon U-test for α=.025 (upper line) and α=.005 (lower line) (after Owen, 1962, from Stange and Henning, 1966, with kind permission of Springer Verlag, Berlin)

n_2 \ n_1	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	–	–	–	–	0	0	0	1	1	1	2	2	2	2	3	3
4	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	–	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
	7	10	13	17	20	24	27	31	34	38	42	45	49	53	57	60
14	13	17	22	26	31	36	40	45	50	55	59	64	69	74	78	83
	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	17	22	28	34	39	45	51	57	63	69	75	81	87	93	99	105
	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
	12	17	22	28	33	39	45	51	57	63	69	74	81	87	93	99
20	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127
	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

Table A-9

Acceptance intervals r_1, r_2 for the Run test r; if only one limit exists, for the other a blank is left ($n_1 \leq n_2$, $n_2 \leq 20$)
(after Swed and Eisenhart, 1943, from Diem, 1962, with kind permission of J.R. Geigy, Basel, and Institute of Mathematical Statistics, Hayward, CA, USA)

n_2 \ α	.05	.01
$n_1 = 2$		
12-20	2-	
$n_1 = 3$		
6-11	2-	
12-14	2-	2-
15-20	2-	2-
$n_1 = 4$		
5, 6	2- 9	
7	2-	
8-14	3-	2-
15	3-	3-
16-20	4-	3-
$n_1 = 5$		
5	2-10	
6	3-10	2-
7	3-11	2-11
8	3-11	2-
9	3-	2-
10	3-	3-
11-17	4-	3-
18-20	5-	4-
$n_1 = 6$		
6	3-11	2-12
7	3-12	2-13
8	3-12	3-13
9-12	4-13	3-
13	5-	3-
14-18	5-	4-
19, 20	6-	4-

n_2 \ α	.05	.01
$n_1 = 7$		
7	3-13	3-13
8	4-13	3-14
9	4-14	3-15
10	5-14	3-15
11	5-14	4-15
12	5-14	4-
13, 14	5-15	4-
15	6-15	4-
16-20	6-	5-
$n_1 = 8$		
8	4-14	3-15
9	5-14	3-15
10, 11	5-15	4-16
12	6-16	4-17
13, 14	6-16	5-17
15	6-16	5-
16	6-17	5-
17	7-17	5-
18-20	7-17	6-
$n_1 = 9$		
9	5-15	4-16
10	5-16	4-17
11	6-16	5-17
12	6-16	5-18
13	6-17	5-18
14	7-17	5-18
15-17	7-18	6-19
18, 19	8-18	6-
20	8-18	7-
$n_1 = 10$		
10	6-16	5-17
11	6-17	5-18
12	7-17	5-19
13	7-18	5-19
14	7-18	6-19
15	7-18	6-20
16	8-19	6-20
17	8-19	7-20
18	8-19	7-21
19	8-20	7-21
20	9-20	7-21

Table A-9 (continued)

n_2 \ α	.05	.01
n_1 = 11		
11	7-17	5-19
12	7-18	6-19
13	7-19	6-20
14	8-19	6-20
15	8-19	7-21
16	8-20	7-21
17, 18	9-20	7-22
19, 20	9-21	8-22
n_1 = 12		
12	7-19	6-20
13	8-19	6-21
14	8-20	7-21
15	8-20	7-22
16	9-21	7-22
17	9-21	8-22
18	9-21	8-23
19, 20	10-22	8-23
n_1 = 13		
13	8-20	7-21
14	9-20	7-22
15	9-21	7-22
16	9-21	8-23
17	10-22	8-23
18	10-22	8-24
19, 20	10-23	9-24
n_1 = 14		
14	9-21	7-23
15	9-22	8-23
16	10-22	8-24
17	10-23	8-24
18	10-23	9-25
19	11-23	9-25
20	11-24	9-25

n_2 \ α	.05	.01
n_1 = 15		
15	10-22	8-24
16	10-23	9-24
17	11-23	9-25
18	11-24	9-25
19	11-24	10-26
20	12-25	10-26
n_1 = 16		
16	11-23	9-25
17	11-24	9-26
18	11-25	10-26
19, 20	12-25	10-27
n_1 = 17		
17	11-25	10-26
18	12-25	10-27
19	12-26	10-27
20	13-26	11-28
n_1 = 18		
18	12-26	11-27
19	13-26	11-28
20	13-27	11-29
n_1 = 19		
19	13-27	11-29
20	13-27	12-29
n_1 = 20		
20	14-28	12-30

Table A-10

Critical values of Wilcoxon's pair difference test
(after Wilcoxon and Wilcox, 1964, from Stange and Henning, 1966,
with kind permission of Springer-Verlag, Berlin)

n \ α	.05	.025	.01	.005
6	2	1	-	-
7	4	2	0	-
8	6	4	2	0
9	8	6	3	2
10	11	8	5	3
11	14	11	7	5
12	17	14	10	7
13	21	17	13	10
14	26	21	16	13
15	30	25	20	16
16	36	30	24	19
17	41	35	28	23
18	47	40	33	28
19	54	46	38	32
20	60	52	43	37
21	68	59	49	43
22	75	66	56	49
23	83	73	62	55
24	92	81	69	61
25	101	90	77	68
30	152	137	120	109
35	214	195	174	160

Table A-11

Critical values of the test $\lambda_n = \frac{\bar{x} - \mu}{R}$

(from Lord, 1947, with kind permission of Biometrika Trustees, London)

n \ α	.05	.025	.01	.005
2	3.196	6.353	15.910	31.828
3	.885	1.304	2.111	3.008
4	.529	.717	1.023	1.316
5	.388	.507	.685	.843
6	.312	.399	.523	.628
7	.263	.333	.429	.507
8	.230	.288	.366	.429
9	.205	.255	.322	.374
10	.186	.230	.288	.333
11	.170	.210	.262	.302
12	.158	.194	.241	.277
13	.147	.181	.224	.256
14	.138	.170	.209	.239
15	.131	.160	.197	.224
16	.124	.151	.186	.212
17	.118	.144	.177	.201
18	.113	.137	.168	.191
19	.108	.131	.161	.182
20	.104	.126	.154	.175

Table A-12

Critical values of the studentized range $Q_{\nu,k}$
(from Harter, 1960, with kind permission of the Institute of Mathematical Statistics, Hayward, CA, USA)

$\alpha = .05$

	2	3	4	5	6	7	8	9	10
1	17.97	26.98	32.82	37.08	40.41	43.12	45.50	47.36	49.07
2	6.085	8.331	9.798	10.88	11.74	12.44	13.03	13.54	13.99
3	4.501	5.910	6.825	7.502	8.037	8.478	8.853	9.177	9.462
4	3.927	5.040	5.757	6.287	6.707	7.053	7.347	7.602	7.826
5	3.635	4.602	5.218	5.637	6.033	6.330	6.582	6.802	6.995
6	3.461	4.339	4.896	5.305	5.628	5.895	6.122	6.319	6.439
7	3.344	4.165	4.681	5.060	5.359	5.606	5.815	5.998	6.158
8	3.261	4.041	4.529	4.886	5.167	5.399	5.597	5.767	5.918
9	3.199	3.949	4.415	4.756	5.024	5.244	5.432	5.595	5.739
10	3.151	3.877	4.327	4.654	4.912	5.124	5.305	5.461	5.599
11	3.113	3.820	4.256	4.574	4.823	5.028	5.202	5.353	5.487
12	3.082	3.773	4.199	4.508	4.751	4.950	5.119	5.265	5.395
13	3.055	3.735	4.151	4.453	4.690	4.885	5.049	5.192	5.318
14	3.033	3.702	4.111	4.407	4.639	4.829	4.990	5.131	5.254
15	3.014	3.674	4.076	4.367	4.595	4.782	4.940	5.077	5.198
16	2.998	3.649	4.046	4.333	4.557	4.741	4.897	5.031	5.150
17	2.984	3.628	4.020	4.303	4.524	4.705	4.858	4.991	5.108
18	2.971	3.609	3.997	4.277	4.495	4.673	4.824	4.956	5.071
19	2.960	3.593	3.977	4.253	4.469	4.645	4.794	4.924	5.038
20	2.950	3.578	3.958	4.232	4.445	4.620	4.768	4.896	5.008
24	2.919	3.532	3.901	4.166	4.373	4.541	4.684	4.807	4.915
30	2.888	3.486	3.845	4.102	4.302	4.464	4.602	4.702	4.824
40	2.858	3.442	3.791	4.039	4.232	4.389	4.521	4.635	4.735
60	2.829	3.399	3.737	3.977	4.163	4.314	4.441	4.550	4.646
120	2.800	3.356	3.685	3.917	4.096	4.241	4.363	4.468	4.560
∞	2.772	3.314	3.633	3.858	4.030	4.170	4.286	4.387	4.474

$\alpha = .01$

k /	2	3	4	5	6	7	8	9	10
1	90.03	135.0	164.3	185.6	202.2	215.8	227.2	237.0	245.6
2	14.04	19.02	22.29	24.72	26.63	28.20	29.53	30.68	31.69
3	8.261	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69
4	6.512	8.120	9.173	9.958	10.58	11.10	11.55	11.93	12.27
5	5.702	6.976	7.804	8.421	8.913	9.321	9.669	9.972	10.24
6	5.243	6.331	7.033	7.556	7.973	8.318	8.613	8.869	9.097
7	4.949	5.919	6.543	7.005	7.373	7.679	7.939	8.166	8.368
8	4.746	5.635	6.204	6.625	6.960	7.237	7.474	7.681	7.863
9	4.596	5.428	5.957	6.348	6.658	6.915	7.134	7.325	7.495
10	4.482	5.270	5.769	6.136	6.428	6.669	6.875	7.055	7.213
11	4.392	5.146	5.621	5.970	6.247	6.476	6.672	6.842	6.992
12	4.320	5.046	5.502	5.836	6.101	6.321	6.507	6.670	6.814
13	4.260	4.964	5.404	5.727	5.981	6.192	6.372	6.528	6.667
14	4.210	4.895	5.322	5.634	5.881	6.085	6.258	6.409	6.543
15	4.168	4.836	5.252	5.556	5.796	5.994	6.162	6.309	6.439
16	4.131	4.786	5.192	5.489	5.722	5.915	6.079	6.222	6.349
17	4.099	4.742	5.140	5.430	5.659	5.847	6.007	6.147	6.270
18	4.071	4.703	5.094	5.379	5.603	5.788	5.944	6.081	6.201
19	4.046	4.670	5.054	5.334	5.554	5.735	5.889	6.022	6.141
20	4.024	4.639	5.018	5.294	5.510	5.688	5.839	5.970	6.087
24	3.956	4.546	4.907	5.168	5.374	5.542	5.685	5.809	5.919
30	3.889	4.455	4.799	5.048	5.242	5.401	5.536	5.653	5.756
40	3.825	4.367	4.696	4.931	5.114	5.265	5.392	5.502	5.599
60	3.762	4.282	4.595	4.818	4.991	5.133	5.253	5.356	5.447
120	3.702	4.200	4.497	4.709	4.872	5.005	5.118	5.214	5.299
	3.643	4.120	4.403	4.603	4.757	4.882	4.987	5.078	5.157

Table A-12 (continued)

$\alpha = .05$

ν \ k	11	12	13	14	15	16	17	18	19
1	50.59	51.96	53.20	54.33	55.36	56.32	57.22	58.04	58.83
2	14.39	14.75	15.08	15.38	15.65	15.91	16.14	16.37	16.57
3	9.717	9.946	10.15	10.35	10.53	10.69	10.84	10.98	11.11
4	8.027	8.208	8.373	8.525	8.664	8.794	8.914	9.028	9.134
5	7.168	7.324	7.466	7.596	7.717	7.828	7.932	8.030	8.122
6	6.649	6.789	6.917	7.034	7.143	7.244	7.338	7.426	7.508
7	6.302	6.431	6.550	6.658	6.759	6.852	6.939	7.020	7.097
8	6.054	6.175	6.287	6.389	6.483	6.571	6.653	6.729	6.802
9	5.867	5.983	6.089	6.186	6.276	6.359	6.437	6.510	6.579
10	5.722	5.833	5.935	6.028	6.114	6.194	6.269	6.339	6.405
11	5.605	5.713	5.811	5.901	5.984	6.062	6.134	6.202	6.265
12	5.511	5.615	5.710	5.798	5.878	5.953	6.023	6.089	6.151
13	5.431	5.533	5.625	5.711	5.789	5.862	5.931	5.995	6.055
14	5.364	5.463	5.554	5.637	5.714	5.786	5.852	5.915	5.974
15	5.306	5.404	5.493	5.574	5.649	5.720	5.785	5.846	5.904
16	5.256	5.352	5.439	5.520	5.593	5.662	5.727	5.786	5.843
17	5.212	5.307	5.392	5.471	5.544	5.612	5.675	5.734	5.790
18	5.174	5.267	5.352	5.429	5.501	5.568	5.630	5.688	5.743
19	5.140	5.231	5.315	5.391	5.462	5.528	5.589	5.647	5.701
20	5.108	5.199	5.282	5.357	5.427	5.493	5.553	5.610	5.663
24	5.012	5.099	5.179	5.251	5.319	5.381	5.439	5.494	5.545
30	4.917	5.001	5.077	5.147	5.211	5.271	5.327	5.379	5.429
40	4.824	4.904	4.977	5.044	5.106	5.163	5.216	5.266	5.313
60	4.732	4.808	4.878	4.942	5.001	5.056	5.107	5.154	5.199
120	4.641	4.714	4.781	4.842	4.898	4.950	4.998	5.044	5.086
∞	4.552	4.622	4.685	4.743	4.796	4.845	4.891	4.934	4.974

$\alpha = .01$

ν \ k	11	12	13	14	15	16	17	18	19
1	253.2	260.0	266.2	271.8	277.0	281.8	286.3	290.4	294.3
2	32.59	33.40	34.13	34.81	35.43	36.00	36.53	37.03	37.50
3	17.13	17.53	17.89	18.22	18.52	18.81	19.07	19.32	19.55
4	12.57	12.84	13.09	13.32	13.53	13.73	13.91	14.08	14.24
5	10.48	10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81
6	9.301	9.485	9.653	9.808	9.951	10.08	10.21	10.32	10.43
7	8.548	8.711	8.860	8.997	9.124	9.242	9.353	9.456	9.554
8	8.027	8.176	8.312	8.436	8.552	8.659	8.760	8.854	8.943
9	7.647	7.784	7.910	8.025	8.132	8.232	8.325	8.412	8.495
10	7.356	7.485	7.603	7.712	7.812	7.906	7.993	8.076	8.153
11	7.128	7.250	7.362	7.465	7.560	7.649	7.732	7.809	7.883
12	6.943	7.060	7.167	7.265	7.356	7.441	7.520	7.594	7.665
13	6.791	6.903	7.006	7.101	7.188	7.269	7.345	7.417	7.485
14	6.664	6.772	6.871	6.962	7.047	7.126	7.199	7.268	7.333
15	6.555	6.660	6.757	6.845	6.927	7.003	7.074	7.142	7.204
16	6.462	6.564	6.658	6.744	6.823	6.898	6.967	7.032	7.093
17	6.381	6.480	6.572	6.656	6.734	6.806	6.873	6.937	6.997
18	6.310	6.407	6.497	6.579	6.655	6.725	6.792	6.854	6.912
19	6.247	6.342	6.430	6.510	6.585	6.654	6.719	6.780	6.837
20	6.191	6.285	6.371	6.450	6.523	6.591	6.654	6.714	6.771
24	6.017	6.106	6.186	6.261	6.330	6.394	6.453	6.510	6.563
30	5.849	5.932	6.008	6.078	6.143	6.203	6.259	6.311	6.361
40	5.686	5.764	5.835	5.900	5.961	6.017	6.069	6.119	6.165
60	5.528	5.601	5.667	5.728	5.785	5.837	5.886	5.931	5.974
120	5.375	5.443	5.505	5.562	5.614	5.662	5.708	5.750	5.790
∞	5.227	5.290	5.348	5.400	5.448	5.493	5.535	5.574	5.611

Table A-12 (continued) $\alpha = .05$

ν \ k	20	22	24	26	28	30	32	34	36
1	59.56	60.91	62.12	63.22	64.23	65.15	66.01	66.81	67.56
2	16.77	17.13	17.45	17.75	18.02	18.27	18.50	18.72	18.92
3	11.24	11.47	11.68	11.87	12.05	12.21	12.36	12.50	12.63
4	9.233	9.418	9.548	9.736	9.875	10.00	10.12	12.23	10.34
5	8.208	8.368	8.512	8.643	8.764	8.875	8.979	9.075	9.165
6	7.587	7.730	7.861	7.979	8.088	8.189	8.283	8.370	8.452
7	7.170	7.303	7.423	7.533	7.634	7.728	7.814	7.895	7.972
8	6.870	6.995	7.109	7.212	7.307	7.395	7.477	7.554	7.625
9	6.644	6.763	6.871	6.970	7.061	7.145	7.222	7.295	7.363
10	6.467	6.582	6.686	6.781	6.868	6.948	7.023	7.093	7.159
11	6.326	6.436	6.536	6.628	6.712	6.790	6.863	6.930	6.994
12	6.209	6.317	6.414	6.503	6.585	6.660	6.731	6.796	6.858
13	6.112	6.217	6.312	6.398	6.478	6.551	6.620	6.684	6.744
14	6.029	6.132	6.224	6.309	6.387	6.459	6.526	6.588	6.647
15	5.958	6.059	6.149	6.233	6.309	6.379	6.445	6.506	6.564
16	5.897	5.995	6.084	6.166	6.241	6.310	6.374	6.434	6.491
17	5.842	5.940	6.027	6.107	6.181	6.249	6.313	6.372	6.427
18	5.794	5.890	5.977	6.055	6.128	6.195	6.258	6.316	6.371
19	5.752	5.846	5.932	6.009	6.081	6.147	6.209	6.267	6.321
20	5.714	5.807	5.891	5.968	6.039	6.104	6.165	6.222	6.275
24	5.594	5.683	5.764	5.838	5.906	5.968	6.027	6.081	6.132
30	5.475	5.561	5.638	5.709	5.774	5.833	5.889	5.941	5.990
40	5.358	5.439	5.513	5.581	5.642	5.700	5.753	5.803	5.849
60	5.241	5.319	5.389	5.453	5.512	5.566	5.617	5.664	5.708
120	5.126	5.200	5.266	5.327	5.382	5.434	5.481	5.526	5.568
	5.012	5.081	5.144	5.201	5.253	5.301	5.346	5.388	5.427

$\alpha = .01$

	20	22	24	26	28	30	32	34	36
1	298.0	304.7	310.8	316.3	321.3	326.0	330.3	334.3	338.0
2	37.95	38.76	39.49	40.15	40.76	41.32	41.84	42.33	42.78
3	19.77	20.17	20.53	20.86	21.16	21.44	21.70	21.95	22.17
4	14.40	14.68	14.93	15.16	15.37	15.57	15.75	15.92	16.08
5	11.93	12.16	12.36	12.54	12.71	12.87	13.02	13.15	13.28
6	10.54	10.73	10.91	11.06	11.21	11.34	11.47	11.58	11.69
7	9.646	9.815	9.970	10.11	10.24	10.36	10.47	10.58	10.67
8	9.027	9.182	8.322	9.450	9.569	9.678	9.779	9.874	9.964
9	8.573	8.717	8.847	8.966	9.075	9.177	9.271	9.360	9.443
10	8.226	8.361	8.483	8.595	8.698	8.794	8.883	8.966	9.044
11	7.952	8.080	8.196	8.303	8.400	8.491	8.575	8.654	8.728
12	7.731	7.853	7.964	8.066	8.159	8.246	8.327	8.402	8.473
13	7.548	7.665	7.772	7.870	7.960	8.043	8.121	8.193	8.262
14	7.395	7.508	7.611	7.705	7.792	7.873	7.948	8.018	8.084
15	7.264	7.374	7.474	7.566	7.650	7.728	7.800	7.869	7.932
16	7.152	7.258	7.356	7.455	7.527	7.602	7.673	7.739	7.802
17	7.053	7.158	7.253	7.340	7.420	7.493	7.563	7.627	7.687
18	6.968	7.070	7.163	7.247	7.325	7.398	7.465	7.528	7.587
19	6.891	6.992	7.082	7.166	7.242	7.313	7.379	7.440	7.498
20	6.823	6.922	7.011	7.092	7.168	7.237	7.302	7.362	7.419
24	6.612	6.705	6.789	6.865	6.936	7.001	7.062	7.119	7.173
30	6.407	6.494	6.572	6.644	6.710	6.772	6.828	6.881	6.932
40	6.209	6.289	6.362	6.429	6.490	6.547	6.600	6.650	6.697
60	6.015	6.090	6.158	6.220	6.227	6.330	6.378	6.424	6.467
120	5.827	5.897	5.959	6.016	6.069	6.117	6.162	6.204	6.244
	5.645	5.709	5.766	5.818	5.866	5.911	5.952	5.990	6.026

Table A-12 (continued)

.05

ν \ k	38	40	50	60	70	80	90	100
1	68.26	68.92	71.73	73.97	75.82	77.40	78.77	79.98
2	19.11	19.28	20.05	20.66	21.16	21.59	21.96	22.29
3	12.75	12.87	13.36	13.76	14.08	14.36	14.61	14.82
4	10.44	10.53	10.93	11.24	11.51	11.73	11.92	12.09
5	9.250	9.330	9.674	9.949	10.18	10.38	10.54	10.69
6	8.529	8.601	8.913	9.163	9.370	9.548	9.702	9.839
7	8.043	8.110	8.400	8.632	8.824	8.989	9.133	9.261
8	7.693	7.756	8.029	8.248	8.430	8.586	8.722	8.843
9	7.428	7.488	7.749	7.958	8.132	8.281	8.410	8.526
10	7.220	7.279	7.529	7.730	7.897	8.041	8.166	8.276
11	7.053	7.110	7.352	7.546	7.708	7.847	7.968	8.075
12	6.916	6.970	7.205	7.394	7.552	7.687	7.804	7.909
13	6.800	6.854	7.083	7.267	7.421	7.552	7.667	7.769
14	6.702	6.754	6.979	7.159	7.309	7.438	7.550	7.650
15	6.618	6.669	6.888	7.065	7.212	7.339	7.449	7.546
16	6.544	6.594	6.810	6.984	7.128	7.252	7.360	7.457
17	6.479	6.529	6.741	6.912	7.054	7.176	7.283	7.377
18	6.422	6.471	6.680	6.848	6.989	7.109	7.213	7.307
19	6.371	6.419	6.626	6.792	6.930	7.048	7.152	7.244
20	6.325	6.373	6.576	6.740	6.877	6.994	7.097	7.187
24	6.181	6.226	6.421	6.579	6.710	6.822	6.920	7.008
30	6.037	6.080	6.267	6.417	6.543	6.650	6.744	6.827
40	5.893	5.934	6.112	6.255	6.375	6.477	6.566	6.645
60	5.750	5.789	5.958	6.093	6.206	6.303	6.387	6.462
120	5.607	5.644	5.802	5.929	6.035	6.126	6.205	6.275
	5.463	5.498	5.646	5.764	5.863	5.947	6.020	6.085

α = .01

	38	40	50	60	70	80	90	100
1	341.5	344.8	358.9	370.1	379.4	387.3	394.1	400.1
2	43.21	43.61	45.33	46.70	47.83	48.80	49.64	50.38
3	22.39	22.59	23.45	24.13	24.71	25.19	25.62	25.99
4	16.23	16.37	16.98	17.46	17.86	18.20	18.50	18.77
5	13.40	13.52	14.00	14.39	14.72	14.99	15.23	15.45
6	11.80	11.90	12.31	12.65	12.92	13.16	13.37	13.55
7	10.77	10.85	11.23	11.52	11.77	11.99	12.17	12.34
8	10.05	10.13	10.47	10.75	10.97	11.17	11.34	11.49
9	9.521	9.594	9.912	10.17	10.38	10.57	10.73	10.87
10	9.117	9.187	9.486	9.726	9.927	10.10	10.25	10.39
11	8.798	8.864	9.148	9.377	9.568	9.732	9.875	10.00
12	8.539	8.603	8.875	9.094	9.277	9.434	9.571	9.693
13	8.326	8.387	8.648	8.859	9.035	9.187	9.318	9.436
14	8.146	8.204	8.457	8.661	8.832	8.978	9.106	9.219
15	7.992	8.049	8.295	8.492	8.658	8.800	8.924	9.035
16	7.860	7.916	8.154	8.347	8.507	8.646	8.767	8.874
17	7.745	7.799	8.031	8.219	8.377	8.511	8.630	8.735
18	7.643	7.696	7.924	8.107	8.261	8.393	8.508	8.611
19	7.553	7.605	7.828	8.008	8.159	8.288	8.401	8.502
20	4.473	7.523	7.742	7.919	8.067	8,194	8.305	8.404
24	7.223	7.270	7.476	7,642	7.780	7.900	8.004	8.097
30	6.978	7.023	7.215	7.370	7.500	7.611	7.709	7.796
40	6.740	6.782	6.960	7.104	7.225	7.328	7.419	7.500
60	6.507	6.546	6.710	6.843	6.954	7.050	7.133	7.207
120	6.281	6.316	6.467	6.588	6.689	6.776	6.852	6.919
	6.060	6.092	6.228	6.338	6.429	6.507	6.575	6.636

Table A-13

Critical values of F_{max} for $\alpha = .05$ (upper line) and $\alpha = .01$ (lower line)
(from David, 1952, with kind permission of Biometrika Trustees, London)

r \ c	2	3	4	5	6	7	8	9	10	11	12
2	39.0	87.5	142.	202.	266.	333.	403.	475.	550.	626.	704.
	199.	448.	729.	1036	1362	1705	2063	2432	2813	3204	3605
3	15.4	27.8	39.2	50.7	62.0	72.9	83.5	93.9	104.	114.	124.
	47.5	85.	120.	151.	184.	21(6)	24(9)	28(1)	31(0)	33(7)	36(1)
4	9.60	15.5	20.6	25.2	29.5	33.6	37.5	41.1	44.6	48.0	51.4
	23.2	37.	49.	59.	69.	79.	89.	97.	106.	113.	120.
5	7.15	10.8	13.7	16.3	18.7	20.8	22.9	24.7	26.5	28.2	29.9
	14.9	22.	28.	33.	38.	42.	46.	50.	54.	57.	60.
6	5.82	8.38	10.4	12.1	13.7	15.0	16.3	17.5	18.6	19.7	20.7
	11.1	15.5	19.1	22.	25.	27.	30.	32.	34.	36.	37.
7	4.99	6.94	8.44	9.70	10.8	11.8	12.7	13.5	14.3	15.1	15.8
	8.89	12.1	14.5	16.5	18.4	20.	22.	23.	24.	26.	27.
8	4.43	6.00	7.18	8.12	9.03	9.78	10.5	11.1	11.7	12.2	12.7
	7.50	9.9	11.7	13.2	14.5	15.8	16.9	17.9	18.9	19.8	21.
9	4.03	5.34	6.31	7.11	7.80	8.41	8.95	9.45	9.91	10.3	10.7
	6.54	8.5	9.9	11.1	12.1	13.1	13.9	14.7	15.3	16.0	16.6
10	3.72	4.85	5.67	6.34	6.92	7.42	7.87	8.28	8.66	9.01	9.34
	5.85	7.4	8.6	9.6	10.4	11.1	11.8	12.4	12.9	13.4	13.9
12	3.28	4.16	4.79	5.30	5.72	6.09	6.42	6.72	7.00	7.25	7.48
	4.91	6.1	6.9	7.6	8.2	8.7	9.1	9.5	9.9	10.2	10.6
15	2.86	3.54	4.01	4.37	4.68	4.95	5.19	5.40	5.59	5.77	5.93
	4.07	4.9	5.5	6.0	6.4	6.7	7.1	7.3	7.5	7.8	8.0
20	2.46	2.95	3.29	3.54	3.76	3.94	4.10	4.24	4.37	4.49	4.59
	3.32	3.8	4.3	4.6	4.9	5.1	5.3	5.5	5.6	5.8	5.9
30	2.07	2.40	2.61	2.78	2.91	3.02	3.12	3.21	3.29	3.36	3.39
	2.63	3.0	3.3	3.4	3.6	3.7	3.8	3.9	4.0	4.1	4.2
60	1.67	1.85	1.96	2.04	2.11	2.17	2.22	2.26	2.30	2.33	2.36
	1.96	2.2	2.3	2.4	2.4	2.5	2.5	2.6	2.6	2.7	2.7
∞	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	1.00	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table A-14

Critical values of Spearman's and Kendall's rank correlation coefficients
(from Stange and Henning, 1968, with kind permission of Springer Verlag, Berlin)

	Spearman's $\lvert r_{s(1-\alpha)} \rvert$			Kendall's $\lvert \tau_{(1-\alpha)} \rvert$		
n \ α	.10	.05	.01	.10	.05	.01
4	1	-	-	1	-	-
5	.90	1	-	.80	1	-
6	.83	.89	1	.73	.87	1
7	.71	.79	.93	.62	.71	.91
8	.64	.74	.88	.57	.64	.79
9	.60	.70	.83	.50	.56	.72
10	.56	.66	.79	.47	.51	.64

Table A-15

Critical values of Dixon's test for outliers r_n
(after Dixon, 1951, from Dunn and Clark, 1974,
with kind permission of Wiley and Sons, New York)

n \ $1-\alpha$	.95	.99	.995
3	.941	.988	.994
4	.765	.889	.926
5	.642	.780	.821
6	.560	.698	.740
7	.507	.637	.680
8	.468	.590	.634
9	.437	.555	.598
10	.412	.527	.568
11	.392	.502	.542
12	.376	.482	.522
13	.361	.465	.503
14	.349	.450	.488
15	.338	.438	.475
16	.329	.426	.463
17	.320	.416	.452
18	.313	.407	.442
19	.306	.398	.433
20	.300	.391	.425
21	.295	.384	.418
22	.290	.378	.411
23	.285	.372	.404
24	.281	.367	.399
25	.277	.362	.393
30	.260	.341	.372

Table A-16

Critical values of Grubbs' test for outliers T_n
(from Grubbs, 1969, with kind permission of
American Statistical Association, Washington, D.C.)

n \ α	.05	.025	.001
3	1.15	1.15	1.15
4	1.46	1.48	1.49
5	1.67	1.71	1.75
6	1.82	1.89	1.94
7	1.94	2.02	2.10
8	2.03	2.13	2.22
9	2.11	2.21	2.32
10	2.18	2.29	2.41
11	2.23	2.36	2.48
12	2.29	2.41	2.55
13	2.33	2.46	2.61
14	2.37	2.51	2.66
15	2.41	2.55	2.71
16	2.44	2.59	2.75
17	2.47	2.62	2.79
18	2.50	2.65	2.82
19	2.53	2.68	2.85
20	2.56	2.71	2.88
21	2.58	2.73	2.91
22	2.60	2.76	2.94
23	2.62	2.78	2.96
24	2.64	2.80	2.99
25	2.66	2.82	3.01
30	2.75	2.91	
50	2.96	3.13	
100	3.21	3.38	

Table A-17

Critical values of $Z_n = \frac{R}{s}$ for outliers
(after David et al., 1954)

n \ $1-\alpha$	.95	.99	.995
	$Z_{(\frac{\alpha}{2})}$ - $Z_{(1-\frac{\alpha}{2})}$	$Z_{(\frac{\alpha}{2})}$ - $Z_{(1-\frac{\alpha}{2})}$	$Z_{(\frac{\alpha}{2})}$ - $Z_{(1-\frac{\alpha}{2})}$
20	3.18 - 4.49	3.01 - 4.79	2.95 - 4.91
30	3.46 - 4.89	3.27 - 5.25	3.22 - 5.39
40	3.66 - 5.15	3.46 - 5.54	3.41 - 5.69
50	3.82 - 5.35	3.61 - 5.77	3.57 - 5.91
60	3.95 - 5.50	3.74 - 5.93	3.69 - 6.09
80	4.15 - 5.73	3.93 - 6.18	3.88 - 6.35
100	4.31 - 5.90	4.09 - 6.36	4.02 - 6.54
150	4.59 - 6.18	4.36 - 6.64	4.30 - 6.84
200	4.78 - 6.38	4.56 - 6.85	4.50 - 7.03
500	5.37 - 6.94	5.13 - 7.42	5.06 - 7.60
1000	5.79 - 7.33	5.57 - 7.80	5.50 - 7.99

Table A-18

Critical values of the mean successive difference test η_n^2 and of the serial correlation coefficient $R_{h.n}$ (underlined values are negative) (from Diem, 1962, with kind permission of J.R. Geigy, Basel)

n	$\eta_n^2(\alpha) - \eta_n^2(1-\alpha)$		$R_{h.n(\alpha)} - R_{h.n(1-\alpha)}$	
	$\alpha = .05$	$\alpha = .01$	$\alpha = .05$	$\alpha = .01$
4	.780 - 3.220	.626 - 3.374		
5	.820 - 3.180	.538 - 3.462	<u>.735</u> - .253	<u>.798</u> - .297
6	.890 - 3.110	.561 - 3.439	<u>.708</u> - .345	<u>.863</u> - .447
7	.936 - 3.064	.614 - 3.386	<u>.674</u> - .370	<u>.799</u> - .510
8	.982 - 3.018	.663 - 3.337	<u>.625</u> - .371	<u>.764</u> - .531
9	1.024 - 2.976	.709 - 3.291	<u>.593</u> - .366	<u>.737</u> - .533
10	1.062 - 2.938	.752 - 3.248	<u>.564</u> - .360	<u>.705</u> - .525
11	1.096 - 2.904	.791 - 3.209	<u>.539</u> - .353	<u>.679</u> - .515
12	1.128 - 2.872	.828 - 3.172	<u>.516</u> - .348	<u>.655</u> - .505
13	1.156 - 2.844	.862 - 3.138	<u>.497</u> - .341	<u>.634</u> - .495
14	1.182 - 2.818	.893 - 3.107	<u>.479</u> - .335	<u>.615</u> - .485
15	1.205 - 2.795	.922 - 3.078	<u>.462</u> - .328	<u>.597</u> - .475
16	1.227 - 2.773	.949 - 3.051	<u>.447</u> - .322	<u>.580</u> - .465
17	1.247 - 2.753	.974 - 3.026	<u>.434</u> - .316	<u>.564</u> - .456
18	1.266 - 2.734	.998 - 3.002	<u>.421</u> - .310	<u>.550</u> - .448
19	1.283 - 2.717	1.020 - 2.980	<u>.410</u> - .304	<u>.536</u> - .440
20	1.300 - 2.700	1.041 - 2.959	<u>.399</u> - .299	<u>.524</u> - .432
21	1.315 - 2.685	1.060 - 2.940	<u>.389</u> - .294	<u>.512</u> - .424
22	1.329 - 2.671	1.078 - 2.922	<u>.380</u> - .289	<u>.502</u> - .417
23	1.342 - 2.658	1.096 - 2.904	<u>.372</u> - .285	<u>.491</u> - .411
24	1.355 - 2.645	1.112 - 2.888	<u>.364</u> - .280	<u>.482</u> - .404
25	1.367 - 2.633	1.128 - 2.872	<u>.356</u> - .276	<u>.473</u> - .398
30	1.418 - 2.582	1.195 - 2.805	<u>.325</u> - .257	<u>.433</u> - .370
35	1.458 - 2.542	1.248 - 2.752	<u>.300</u> - .242	<u>.402</u> - .348
40	1.492 - 2.508	1.293 - 2.707	<u>.279</u> - .229	<u>.377</u> - .330
45	1.521 - 2.479	1.332 - 2.668	<u>.262</u> - .218	<u>.356</u> - .314
50	1.544 - 2.456	1.363 - 2.637	<u>.248</u> - .208	<u>.338</u> - .300
75	1.625 - 2.375	1.471 - 2.529	<u>.199</u> - .173	<u>.276</u> - .250
100	1.674 - 2.326	1.539 - 2.461	<u>.175</u> - .154	<u>.243</u> - .222
200	1.768 - 2.232	1.673 - 2.327	<u>.121</u> - .111	<u>.170</u> - .160
1000			<u>.053</u> - .051	<u>.075</u> - .073
∞	2.000 - 2.000	2.000 - 2.000		

Table A-19

The probit transformation $Y(x_j)$
(from Finney, 1971, with kind permission of C. Griffin, High Wycombe and London)

Cumulated frequencies $F(x_j)$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	-	2.67	2.95	3.12	3.25	3.36	3.45	3.52	3.59	3.66
0.10	3.72	3.77	3.82	3.87	3.92	3.96	4.01	4.05	4.08	4.12
0.20	4.16	4.19	4.23	4.26	4.29	4.33	4.36	4.39	4.42	4.45
0.30	4.48	4.50	4.53	4.56	4.59	4.61	4.64	4.67	4.69	4.72
0.40	4.75	4.77	4.80	4.82	4.85	4.87	4.90	4.92	4.95	4.97
0.50	5.00	5.03	5.05	5.08	5.10	5.13	5.15	5.18	5.20	5.23
0.60	5.25	5.28	5.31	5.33	5.36	5.39	5.41	5.44	5.47	5.50
0.70	5.52	5.55	5.58	5.61	5.64	5.67	5.71	5.74	5.77	5.81
0.80	5.84	5.88	5.92	5.95	5.99	6.04	6.08	6.13	6.18	6.23
0.90	6.28	6.34	6.41	6.48	6.55	6.64	6.75	6.88	7.05	7.33

Cumulated frequencies $F(x_j)$	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.97	6.88	6.90	6.91	6.93	6.94	6.96	6.98	7.00	7.01	7.03
0.98	7.05	7.07	7.10	7.12	7.14	7.17	7.20	7.23	7.26	7.29
0.99	7.33	7.37	7.41	7.46	7.51	7.58	7.65	7.75	7.88	8.09

Table A-20

Minimum and maximum working probits, range, and weighting coefficients
(from Finney, 1971, with kind permission of C. Griffin, High Wycombe and London)

Expected probit $Y'(x_j)$	Minimum working probit Y_0	Range A	Maximum working probit Y_1	Expected probit $Y'(x_j)$	Weighting coefficient w
1.1	0.8579	5 034.	9.1421	8.9	.00082
1.2	0.9522	3 425.	9.0478	8.8	.00118
1.3	1.0462	2 354.	8.9538	8.7	.00167
1.4	1.1400	1 634.	8.8600	8.6	.00235
1.5	1.2334	1 146.	8.7666	8.5	.00327
1.6	1.3266	811.5	8.6734	8.4	.00451
1.7	1.4194	580.5	8.5806	8.3	.00614
1.8	1.5118	419.4	8.4882	8.2	.00828
1.9	1.6038	306.1	8.3962	8.1	.01104
2.0	1.6954	225.6	8.3046	8.0	.01457
2.1	1.7866	168.00	8.2134	7.9	.01903
2.2	1.8772	126.34	8.1228	7.8	.02458
2.3	1.9673	95.96	8.0327	7.7	.03143
2.4	2.0568	73.62	7.9432	7.6	.03977
2.5	2.1457	57.05	7.8543	7.5	.04979
2.6	2.2339	44.654	7.7661	7.4	.06168
2.7	2.3214	35.302	7.6786	7.3	.07564
2.8	2.4081	28.189	7.5919	7.2	.09179
2.9	2.4938	22.736	7.5062	7.1	.11026
3.0	2.5786	18.522	7.4214	7.0	.13112
3.1	2.6624	15.2402	7.3376	6.9	.15436
3.2	2.7449	12.6662	7.2551	6.8	.17994
3.3	2.8261	10.6327	7.1739	6.7	.20774
3.4	2.9060	9.0154	7.0940	6.6	.23753
3.5	2.9842	7.7210	7.0158	6.5	.26907
3.6	3.0606	6.6788	6.9394	6.4	.30199
3.7	3.1351	5.8354	6.8649	6.3	.33589
3.8	3.2074	5.1497	6.7926	6.2	.37031
3.9	3.2773	4.5903	6.7227	6.1	.40474
4.0	3.3443	4.1327	6.6557	6.0	.43863
4.1	3.4083	3.7582	6.5917	5.9	.47144
4.2	3.4687	3.4519	6.5313	5.8	.50260
4.3	3.5251	3.2025	6.4749	5.7	.53159
4.4	3.5770	3.0010	6.4230	5.6	.55788
4.5	3.6236	2.8404	6.3764	5.5	.58099
4.6	3.6643	2.7154	6.3357	5.4	.60052
4.7	3.6982	2.6220	6.3018	5.3	.61609
4.8	3.7241	2.5573	6.2759	5.2	.62742
4.9	3.7407	2.5192	6.2593	5.1	.63431
5.0	3.7467	2.5066	6.2533	5.0	.63662
5.1	3.7401	2.5192	6.2599	4.9	.63431
5.2	3.7187	2.5573	6.2813	4.8	.62742
5.3	3.6798	2.6220	6.3202	4.7	.61609
5.4	3.6203	2.7154	6.3797	4.6	.60052
5.5	3.5360	2.8404	6.4640	4.5	.58099
5.6	3.4220	3.0010	6.5780	4.4	.55788
5.7	3.2724	3.2025	6.7276	4.3	.53159
5.8	3.0794	3.4519	6.9206	4.2	.50260
5.9	2.8335	3.7582	7.1665	4.1	.47144
6.0	2.5229	4.1327	7.4771	4.0	.43863
6.1	2.1325	4.5903	7.8675	3.9	.40474
6.2	1.6429	5.1497	8.3571	3.8	.37031
6.3	1.0295	5.8354	8.9705	3.7	.33589
6.4	0.2606	6.6788	9.7394	3.6	.30199
6.5	-0.7051	7.7210	10.7051	3.5	.26907

Table A-21

Critical values of Van der Waerden's X test
(from Van der Waerden, 1971, with kind permission of Springer Verlag, Berlin)

$\alpha = .025$				$\alpha = .01$				$\alpha = .005$			
n	$\lvert n_1-n_2\rvert =$			n	$\lvert n_1-n_2\rvert =$			n	$\lvert n_1-n_2\rvert =$		
	0,1	2,3	4,5		0,1	2,3	4,5		0,1	2,3	4,5
6	∞	∞	∞	6	∞	∞	∞	6	∞	∞	∞
7	∞	∞	∞	7	∞	∞	∞	7	∞	∞	∞
8	2.40	2.30	∞	8	∞	∞	∞	8	∞	∞	∞
9	2.38	2.20	∞	9	2.80	∞	∞	9	∞	∞	∞
10	2.60	2.49	2.30	10	3.00	2.90	2.80	10	3.20	3.10	∞
11	2.72	2.58	2.40	11	3.20	3.00	2.90	11	2.40	3.40	∞
12	2.86	2.79	2.68	12	3.29	3.30	3.20	12	3.60	3.58	3.40
13	2.96	2.91	2.78	13	3.50	3.36	3.18	13	3.71	3.68	3.50
14	3.11	3.06	3.00	14	3.62	3.55	3.46	14	3.94	3.88	3.76
15	3.24	3.19	3.06	15	3.74	3.68	3.57	15	4.07	4.05	3.88
16	3.39	3.36	3.28	16	3.92	3.90	3.80	16	4.26	4.25	4.12
17	3.49	3.44	3.36	17	4.06	4.01	3.90	17	4.44	4.37	4.23
18	3.63	3.60	3.53	18	4.23	4.21	4.14	18	4.60	4.58	4.50
19	3.73	3.69	3.61	19	4.37	4.32	4.23	19	4.77	4.71	4.62
20	3.86	3.84	3.78	20	4.52	4.50	4.44	20	4.94	4.92	4.85
21	3.96	3.92	3.85	21	4.66	4.62	4.53	21	5.10	5.05	4.96
22	4.08	4.06	4.01	22	4.80	4.78	4.72	22	5.26	5.24	5.17
23	4.18	4.15	4.08	23	4.92	4.89	4.81	23	5.40	5.36	5.27
24	4.29	4.27	4.23	24	5.06	5.04	4.99	24	5.55	5.53	5.48
25	4.39	4.36	4.30	25	5.18	5.14	5.08	25	5.68	5.65	5.58
26	4.50	4.48	4.44	26	5.30	5.29	5.24	26	5.83	5.81	5.76
27	4.59	4.56	4.51	27	5.42	5.39	5.33	27	5.95	5.92	5.85
28	4.69	4.68	4.64	28	5.54	5.52	5.48	28	6.09	6.07	6.03
29	4.78	4.76	4.72	29	5.65	5.62	5.57	29	6.22	6.19	6.13
30	4.88	4.87	4.84	30	5.77	5.75	5.72	30	6.35	6.34	6.30
31	4.97	4.95	4.91	31	5.87	5.85	5.80	31	6.47	6.44	6.39
32	5.07	5.06	5.03	32	5.99	5.97	5.94	32	6.60	6.58	6.55
33	5.15	5.13	5.10	33	6.09	6.07	6.02	33	6.71	6.69	6.64
34	5.25	5.24	5.21	34	6.20	6.19	6.16	34	6.84	6.82	6.79
35	5.33	5.31	5.28	35	6.30	6.28	6.24	35	6.95	6.92	6.88
36	5.42	5.41	5.38	36	6.40	6.39	6.37	36	7.06	7.05	7.02
37	5.50	5.48	5.45	37	6.50	6.48	6.45	37	7.17	7.15	7.11
38	5.59	5.58	5.55	38	6.60	6.59	6.57	38	7.28	7.27	7.25
39	5.67	5.65	5.62	39	6.70	6.68	6.65	39	7.39	7.37	7.33
40	5.75	5.74	5.72	40	6.80	6.79	6.77	40	7.50	7.49	7.47
41	5.83	5.81	5.79	41	6.89	6.88	6.85	41	7.62	7.60	7.56
42	5.91	5.90	5.88	42	6.99	6.98	6.96	42	7.72	7.71	7.69
43	5.99	5.97	5.95	43	7.08	7.07	7.04	43	7.82	7.81	7.77
44	6.06	6.06	6.04	44	7.17	7.17	7.14	44	7.93	7.92	7.90
45	6.14	6.12	6.10	45	7.26	7.25	7.22	45	8.02	8.01	7.98
46	6.21	6.21	6.19	46	7.35	7.35	7.32	46	8.13	8.12	8.10
47	6.29	6.27	6.25	47	7.44	7.43	7.40	47	8.22	8.21	8.18
48	6.36	6.35	6.34	48	7.53	7.52	7.50	48	8.32	8.31	8.29
49	6.43	6.42	6.39	49	7.61	7.60	7.57	49	8.41	8.40	8.37
50	6.50	6.50	6.48	50	7.70	7.69	7.68	50	8.51	8.50	8.48

Table A-22

Critical lower values $L_{(\alpha)}$ of the binomial distribution
(from Diem, 1962, with kind permission of J.R. Geigy, Basel)

n	$L_{(.01)}$	$L_{(.05)}$	n	$L_{(.01)}$	$L_{(.05)}$
6	-	0	41	11	13
7	-	0	42	12	14
8	0	0	43	12	14
9	0	1	44	13	15
10	0	1	45	13	15
11	0	1	46	13	15
12	1	2	47	14	16
13	1	2	48	14	16
14	1	2	49	15	17
15	2	3	50	15	17
16	2	3	51	15	18
17	2	4	52	16	18
18	3	4	53	16	18
19	3	4	54	17	19
20	3	5	55	17	19
21	4	5	56	17	20
22	4	5	57	18	20
23	4	6	58	18	21
24	5	6	59	19	21
25	5	7	60	19	21
26	6	7	61	20	22
27	6	7	62	20	22
28	6	8	63	20	23
29	7	8	64	21	23
30	7	9	65	21	24
31	7	9	66	22	24
32	8	9	67	22	25
33	8	10	68	22	25
34	9	10	69	23	25
35	9	11	70	23	26
36	9	11	80	28	30
37	10	12	90	32	35
38	10	12	100	36	39
39	11	12	200	81	85
40	11	13	500	220	227

Table A-23

Coefficients a_i for the W test for normality
(from Shapiro and Wilk, 1965, with kind permission of Biometrika Trustees, London)

i \ n		2	3	4	5	6	7	8	9	10
1		0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739
2		-	.0000	.1677	.2413	.2806	.3031	.3164	.3244	.3291
3		-	-	-	.0000	.0875	.1401	.1743	.1976	.2141
4		-	-	-	-	-	.0000	.0561	.0947	.1224
5		-	-	-	-	-	-	-	.0000	.0399

i \ n	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	.3315	.3325	.3325	.3318	.3306	.3290	.3273	.3253	.3232	.3211
3	.2260	.2347	.2412	.2460	.2495	.2521	.2540	.2553	.2561	.2565
4	.1429	.1586	.1707	.1802	.1878	.1939	.1988	.2027	.2059	.2085
5	.0695	.0922	.1099	.1240	.1353	.1447	.1524	.1587	.1641	.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	-	-	.0000	.0240	.0433	.0593	.0752	.0837	.0932	.1013
8	-	-	-	-	.0000	.0196	.0359	.0496	.0612	.0711
9	-	-	-	-	-	-	.0000	.0163	.0303	.0422
10	-	-	-	-	-	-	-	-	.0000	.0140

i \ n	21	22	23	24	25	26	27	28	29	30
1	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407	0.4366	0.4328	0.4291	0.4254
2	.3185	.3156	.3126	.3098	.3069	.3043	.3018	.2992	.2968	.2944
3	.2578	.2571	.2563	.2554	.2543	.2533	.2522	.2510	.2499	.2487
4	.2119	.2131	.2139	.2145	.2148	.2151	.2152	.2151	.2150	.2148
5	.1736	.1764	.1787	.1807	.1822	.1836	.1848	.1857	.1864	.1870
6	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563	0.1584	0.1601	0.1616	0.1630
7	.1092	.1150	.1201	.1245	.1283	.1316	.1346	.1372	.1395	.1415
8	.0804	.0878	.0941	.0997	.1046	.1089	.1128	.1162	.1192	.1219
9	.0530	.0618	.0696	.0764	.0823	.0876	.0923	.0965	.1002	.1036
10	.0263	.0368	.0459	.0539	.0610	.0672	.0728	.0778	.0822	.0862
11	0.0000	0.0122	0.0228	0.0321	0.0403	0.0476	0.0540	0.0598	0.0650	0.0697
12	-	-	.0000	.0107	.0200	.0284	.0358	.0424	.0483	.0537
13	-	-	-	-	.0000	.0094	.0178	.0253	.0320	.0381
14	-	-	-	-	-	-	.0000	.0084	.0159	.0227
15	-	-	-	-	-	-	-	-	.0000	.0076

Continued

Table A-23 (continued)

i \ n	31	32	33	34	35	36	37	38	39	40
1	0.4220	0.4188	0.4156	0.4127	0.4096	0.4068	0.4040	0.4015	0.3989	0.3964
2	.2921	.2898	.2876	.2854	.2834	.2813	.2794	.2774	.2755	.2737
3	.2475	.2463	.2451	.2439	.2427	.2415	.2403	.2391	.2380	.2368
4	.2145	.2141	.2137	.2132	.2127	.2121	.2116	.2110	.2104	.2098
5	.1874	.1878	.1880	.1882	.1883	.1883	.1883	.1881	.1880	.1878
6	0.1641	0.1651	0.1660	0.1667	0.1673	0.1678	0.1683	0.1686	0.1689	0.1691
7	.1433	.1449	.1463	.1475	.1487	.1496	.1505	.1513	.1520	.1526
8	.1243	.1265	.1284	.1301	.1317	.1331	.1344	.1356	.1366	.1376
9	.1066	.1093	.1118	.1140	.1160	.1179	.1196	.1211	.1225	.1237
10	.0899	.0931	.0961	.0988	.1013	.1036	.1056	.1075	.1092	.1108
11	0.0739	0.0777	0.0812	0.0844	0.0873	0.0900	0.0924	0.0947	0.0967	0.0986
12	.0585	.0629	.0669	.0706	.0739	.0770	.0798	.0824	.0848	.0870
13	.0435	.0485	.0530	.0572	.0610	.0645	.0677	.0706	.0733	.0759
14	.0289	.0344	.0395	.0441	.0484	.0523	.0559	.0592	.0622	.0651
15	.0144	.0206	.0262	.0314	.0361	.0404	.0444	.0481	.0515	.0546
16	0.0000	0.0068	0.0131	0.0187	0.0239	0.0287	0.0331	0.0372	0.0409	0.0444
17	-	-	.0000	.0062	.0119	.0172	.0220	.0264	.0305	.0343
18	-	-	-	-	.0000	.0057	.0110	.0158	.0203	.0244
19	-	-	-	-	-	-	.0000	.0053	.0101	.0146
20	-	-	-	-	-	-	-	-	.0000	.0049

i \ n	41	42	43	44	45	46	47	48	49	50
1	0.3940	0.3917	0.3894	0.3872	0.3850	0.3830	0.3808	0.3789	0.3770	0.3751
2	.2719	.2701	.2684	.2667	.2651	.2635	.2620	.2604	.2589	.2574
3	.2357	.2345	.2334	.2323	.2313	.2302	.2291	.2281	.2271	.2260
4	.2091	.2085	.2078	.2072	.2065	.2058	.2052	.2045	.2038	.2032
5	.1876	.1874	.1871	.1868	.1865	.1862	.1859	.1855	.1851	.1847
6	0.1693	0.1694	0.1695	0.1695	0.1695	0.1695	0.1695	0.1693	0.1692	0.1691
7	.1531	.1535	.1539	.1542	.1545	.1548	.1550	.1551	.1553	.1554
8	.1384	.1392	.1398	.1405	.1410	.1415	.1420	.1423	.1427	.1430
9	.1349	.1259	.1269	.1278	.1286	.1293	.1300	.1306	.1312	.1317
10	.1123	.1136	.1149	.1160	.1170	.1180	.1189	.1197	.1205	.1212
11	0.1004	0.1020	0.1035	0.1049	0.1062	0.1073	0.1085	0.1095	0.1105	0.1113
12	.0891	.0909	.0927	.0943	.0959	.0972	.0986	.0998	.1010	.1020
13	.0782	.0804	.0824	.0842	.0860	.0876	.0892	.0906	.0919	.0932
14	.0677	.0701	.0724	.0745	.0765	.0783	.0801	.0817	.0832	.0846
15	.0575	.0602	.0628	.0651	.0673	.0694	.0713	.0731	.0748	.0764
16	0.0476	0.0506	0.0534	0.0560	0.0584	0.0607	0.0628	0.0648	0.0667	0.0685
17	.0379	.0411	.0442	.0471	.0497	.0522	.0546	.0568	.0588	.0608
18	.0283	.0318	.0352	.0383	.0412	.0439	.0465	.0489	.0511	.0532
19	.0188	.0227	.0263	.0296	.0328	.0357	.0385	.0411	.0436	.0459
20	.0094	.0136	.0175	.0211	.0245	.0277	.0307	.0335	.0361	.0386
21	0.0000	0.0045	0.0087	0.0126	0.0163	0.0197	0.0229	0.0259	0.0288	0.0314
22	-	-	.0000	.0042	.0081	.0118	.0153	.0185	.0215	.0244
23	-	-	-	-	.0000	.0039	.0076	.0111	.0143	.0174
24	-	-	-	-	-	-	.0000	.0037	.0071	.0104
25	-	-	-	-	-	-	-	-	.0000	.0035

Table A-24

Critical values of the W test for normality
(from Shapiro and Wilk, 1965, with kind permission of Biometrika Trustees, London)

n \ α	0.01	0.02	0.05	0.10
3	0.753	0.756	0.767	0.789
4	.687	.707	.748	.792
5	.686	.715	.762	.806
6	0.713	0.743	0.788	0.826
7	.730	.760	.803	.838
8	.749	.778	.818	.851
9	.746	.791	,829	,859
10	.781	.806	.842	.869
11	0.792	0.817	0.850	0.876
12	.805	.828	.859	.883
13	.814	.837	.866	.889
14	.825	.846	.874	.895
15	.835	.855	.881	.901
16	0.844	0.863	0.887	0.906
17	.851	.869	.892	.910
18	.858	.874	.897	.914
19	.863	.879	.901	.917
20	.868	.884	.905	.920
21	0.873	0.888	0.908	0.923
22	.878	.892	.911	.926
23	.881	.895	.914	.928
24	.884	.898	.916	.930
25	.888	.901	.918	.931
26	0.891	0.904	0.920	0.933
27	.894	.906	.923	.935
28	.896	.908	.924	.936
29	.898	.910	.926	.937
30	.900	.912	.927	.939
31	0.902	0.914	0.929	0.940
32	.904	.915	.930	.941
33	.906	.917	.931	.942
34	.908	.919	.933	.943
35	.910	.920	.934	.944
36	0.912	0.922	0.935	0.945
37	.914	.924	.936	.946
38	.916	.925	.938	.947
39	.917	.927	.939	.948
40	.919	.928	.940	.949
41	0.920	0.929	0.941	0.950
42	.922	.930	.942	.951
43	.923	.932	.943	.951
44	.924	.933	.944	.952
45	.926	.934	.945	.953
46	0.927	0.935	0.945	0.935
47	.928	.936	.946	.954
48	.929	.937	.947	.954
49	.929	.937	.947	.955
50	.930	.938	.947	.955

Table A-25

Critical values of the normal probability plot correlation coefficient r_n
(from Filliben, 1975, with kind permission of
American Statistical Association, Washington, D.C.)

n	α					
	.000	.005	.01	.025	.05	.10
3	.866	.867	.869	.872	.879	.891
4	.784	.813	.822	.845	.868	.894
5	.726	.803	.822	.855	.879	.902
6	.683	.818	.835	.868	.890	.911
7	.684	.828	.847	.876	.899	.916
8	.619	.841	.859	.886	.905	.924
9	.595	.851	.868	.893	.912	.929
10	.574	.860	.876	.900	.917	.934
11	.556	.868	.883	.906	.922	.938
12	.539	.875	.889	.912	.926	.941
13	.525	.882	.895	.917	.931	.944
14	.512	.888	.901	.921	.934	.947
15	.500	.894	.907	.925	.937	.950
16	.489	.899	.912	.928	.940	.952
17	.478	.903	.916	.931	.942	.954
18	.469	.907	.919	.934	.945	.956
19	.460	.909	.923	.937	.947	.958
20	.452	.912	.925	.939	.950	.960
21	.445	.914	.928	.942	.952	.961
22	.437	.918	.930	.944	.954	.962
23	.431	.922	.933	.947	.955	.964
24	.424	.926	.936	.949	.957	.965
25	.418	.928	.937	.950	.958	.966
26	.412	.930	.939	.952	.959	.967
27	.407	.932	.941	.953	.960	.968
28	.402	.934	.943	.955	.962	.969
29	.397	.937	.945	.956	.962	.969
30	.392	.938	.947	.957	.964	.970
31	.388	.939	.948	.958	.965	.971
32	.383	.939	.949	.959	.966	.972
33	.379	.940	.950	.960	.967	.973
34	.375	.941	.951	.960	.967	.973
35	.371	.943	.952	.961	.968	.974
36	.367	.945	.953	.962	.968	.974
37	.364	.947	.955	.962	.969	.975
38	.360	.948	.956	.964	.970	.975
39	.357	.949	.957	.965	.971	.976
40	.354	.949	.958	.966	.972	.977
41	.351	.950	.958	.967	.972	.977
42	.348	.951	.959	.967	.973	.978
43	.345	.953	.959	.967	.973	.978
44	.342	.954	.960	.968	.973	.978
45	.339	.955	.961	.969	.974	.978
46	.336	.956	.962	.969	.974	.979
47	.334	.956	.963	.970	.974	.979
48	.331	.957	.963	.970	.975	.980
49	.329	.957	.964	.971	.975	.980
50	.326	.959	.965	.972	.977	.981
55	.315	.962	.967	.974	.978	.982
60	.305	.965	.970	.976	.980	.983
65	.296	.967	.972	.977	.981	.984
70	.288	.969	.974	.978	.982	.985
75	.281	.971	.975	.979	.983	.986
80	.274	.973	.976	.980	.984	.987
85	.268	.974	.977	.981	.985	.987
90	.263	.976	.978	.982	.985	.988
95	.257	.977	.979	.983	.986	.989
100	.252	.979	.981	.984	.987	.989

INTERPOLATION

Tables generally furnish the value of a function z which depends on 1, sometimes 2, argumental (= independent), variables x, y. Not always all values of the argumental variable (= arguments) are recorded so that the correspondent functional values can only be guessed by interpolation.

Linear Interpolation

The argument x for which the function z is searched, is comprised between x_1 and x_2. From the table we obtain z_1 and z_2 in correspondence to x_1 and x_2, then the interpolated value z corresponding to x, is obtained by

$$(*) \qquad z = z_1 + \frac{x-x_1}{x_2-x_1}(z_2-z_1)$$

Example i. We want the critical value of t_{33} and $\alpha = .01$. We get from Table A-4

$x_1 = 30$	$x = 33$	$x_2 = 40$
$t_{30} = z_1 = 2.457$	$t_{33} \approx z = ?$	$t_{40} = z_2 = 2.423$

On using (*) we get

$$z = 2.457 + \frac{3}{10}(.034) = 2.4468$$

Example ii. Having obtained from an experiment $t_{25} = 1.935$ we would like to know the value of α. From Table A-4 we have now

$t_{25(.95)} = x = 1.708$	$t_{25(1-\alpha)} = x = 1.935$	$t_{25(.975)} = x_2 = 2.060$
$\alpha_1 = z_1 = .05$	$\alpha \approx z = ?$	$\alpha_2 = z_2 = .025$

Hence

$$z = .05 + \frac{.227}{.325}(-.025) = .0339$$

Harmonic Interpolation

In some tables the last arguments are very high or even infinite. In these cases a harmonic interpolation is indicated. This is achieved by choosing the last finite argument and dividing it by the other arguments, e.g. as in Table A-6.1. The last finite argument is 120 and so we transform the last arguments to

$$\frac{120}{\infty} = 0, \quad \frac{120}{120} = 1, \quad \frac{120}{60} = 2, \quad \frac{120}{40} = 3, \quad \frac{120}{30} = 4, \quad \text{etc.}$$

Since these values are linear, we can apply to them (*).

Example iii. In Table A-6.1 we look for the value of $F_{2,150}$. We put

$$x_1 = \frac{120}{\infty} = 0 \qquad x = \frac{120}{150} = .8 \qquad x_2 = \frac{120}{120} = 1$$

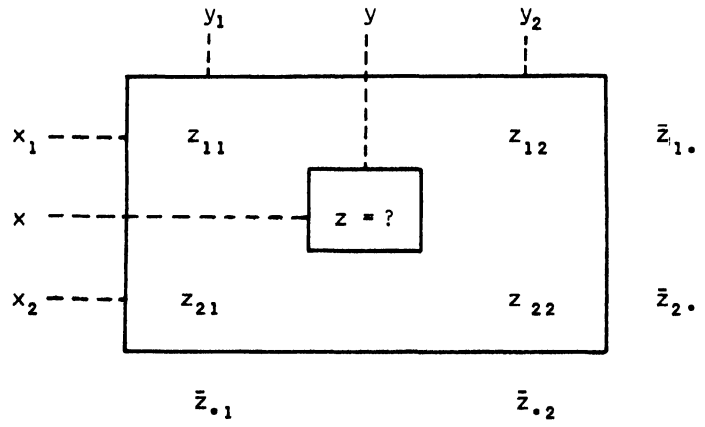
$$z_1 = 3.00 \qquad F_{2,150} \approx z = ? \qquad z_2 = 3.07$$

Applying (*) we obtain

$$z = 3.00 + \frac{.8}{1} \; .07 = 3.056$$

Areal Interpolation

If the function z depends on 2 variables x and y it may be interesting to guess the function of 2 values which are not reported in the table. If we indicate by x_1, x_2, and y_1, y_2 the couples of argumental values which comprise x and y, then we get from the table four values of the function z, for which we evaluate the means for each row and column, thus obtaining

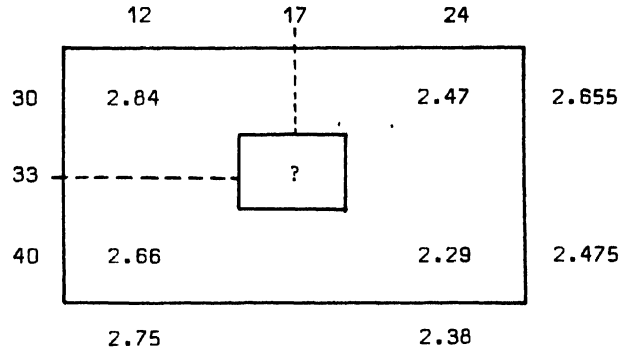


Formula (*) becomes now

$$(**) \qquad z = z_{11} + \frac{x-x_1}{x_2-x_1} \; (\bar{z}_{2.} - \bar{z}_{1.}) + \frac{y-y_1}{y_2-y_1} \; (\bar{z}_{.2} - \bar{z}_{.1})$$

This is equivalent to (*) first applied to x and then to y or *vice versa*, but only if neither x nor y have been transformed for harmonic interpolation.

Example iv. We look for $F_{17,33}$ in Table A-6.3. 17 is comprised between 12 and 24, 33 between 30 and 40. Hence,



By (***) we obtain

$$z = 2.84 + \frac{3}{10}\ (-.18) + \frac{5}{12}\ (-.37) = 2.63 \approx F_{17,33}$$

If we transform x and y harmonically, we get

$$x_1 = \frac{40}{30} = 1.\bar{3} \quad ; \quad x_2 = \frac{40}{40} = 1 \quad ; \quad x = \frac{40}{33} = 1.\overline{21} \quad ;$$

$$y_1 = \frac{24}{12} = 2 \quad ; \quad y_2 = \frac{24}{24} = 1 \quad ; \quad y = \frac{24}{17} = 1.412 \quad ;$$

Hence, applying (***)

$$z = 2.557 \approx F_{17,33}$$

The true value $F_{17,33(.99)}$ lies between 2.557 and 2.63.

NO: 11130

131 + 17

148